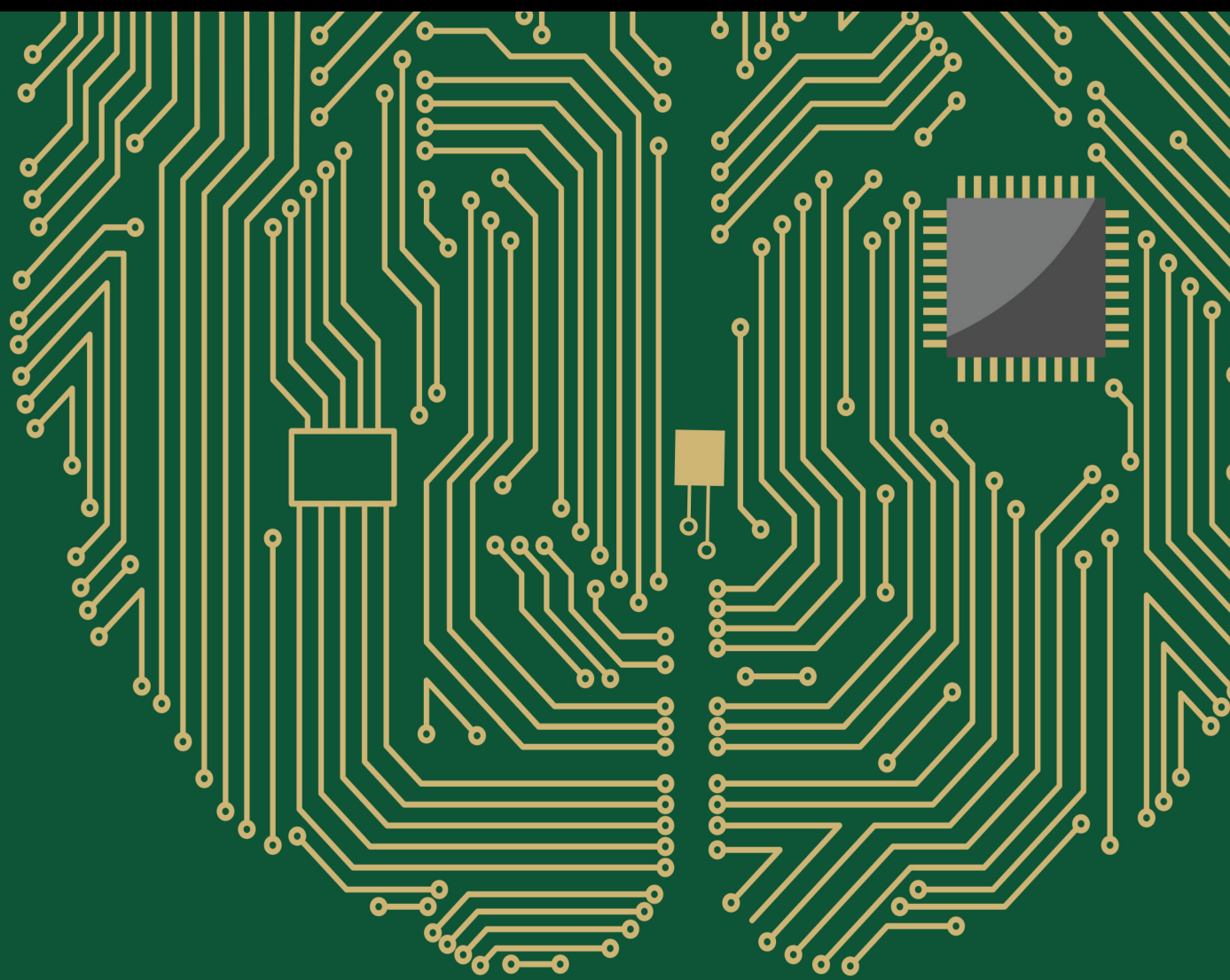# Advances in Computational Intelligence Techniques for Next Generation Internet of Things

Lead Guest Editor: Akshi Kumar
Guest Editors: Sumarga Kumar Sah Tyagi and Danda B. Rawat

Advances in Computational Intelligence Techniques for Next Generation Internet of Things

Advances in Computational Intelligence Techniques for Next Generation Internet of Things

Lead Guest Editor: Akshi Kumar
Guest Editors: Sumarga Kumar Sah Tyagi and
Danda B. Rawat

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Retraction

Retracted: A Social-aware and Mobile Computing-based E-Commerce Product Recommendation System

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Xu and J. Wang, "A Social-aware and Mobile Computing-based E-Commerce Product Recommendation System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9501246, 8 pages, 2022.

Retraction

Retracted: IoT-Oriented Wireless Sensor Network and Sports Dance Movement Perception

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Zhu, "IoT-Oriented Wireless Sensor Network and Sports Dance Movement Perception," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6477170, 8 pages, 2022.

Hindawi

Research Article

IoT-Based Reinforcement Learning Using Probabilistic Model for Determining Extensive Exploration through Computational Intelligence for Next-Generation Techniques

Pradeep Kumar Tiwari [1] **Pooja Singh** [2] **Navaneetha Krishnan Rajagopal,** [3] **K. Deepa,** [4] **Sampada Gulavani,** [5] **Amit Verma,** [6] **and Yekula Prasanna Kumar** [7]

[1]*Birla Global University, Gothapatna, Bhubaneswar, Odisha, India*
[2]*School of Computing Science & Engineering, Department of CSE, Galgotias University, Greater Noida, UP, India*
[3]*Business Studies, The University of Technology and Applied Sciences, Salalah, Oman*
[4]*Department of Computer Science and Engineering, M. Kumarasamy College of Engineering, Thalavapalayam, Karur, Tamilnadu, India*
[5]*Department of MCA, Bharati Vidyapeeth (Deemed to Be University) Institute of Management, Kolhapur, Maharashtra, India*
[6]*University Centre for Research and Development, Department of Computer Science and Engineering, Chandigarh University Gharuan, Mohali, Punjab, India*
[7]*Department of Mining Engineering, College of Engineering and Technology, Bule Hora University, Blue Hora 144, Oromia Region, Ethiopia*

Correspondence should be addressed to Yekula Prasanna Kumar; prasannaky@bhu.edu.et

Computing intelligence is built on several learning and optimization techniques. Incorporating cutting-edge learning techniques to balance the interaction between exploitation and exploration is therefore an inspiring field, especially when it is combined with IoT. The reinforcement learning techniques created in recent years have largely focused on incorporating deep learning technology to improve the generalization skills of the algorithm while ignoring the issue of detecting and taking full advantage of the dilemma. To increase the effectiveness of exploration, a deep reinforcement algorithm based on computational intelligence is proposed in this study, using intelligent sensors and the Bayesian approach. In addition, the technique for computing the posterior distribution of parameters in Bayesian linear regression is expanded to nonlinear models such as artificial neural networks. The Bayesian Bootstrap Deep Q-Network (BBDQN) algorithm is created by combining the bootstrapped DQN with the recommended computing technique. Finally, tests in two scenarios demonstrate that, when faced with severe exploration problems, BBDQN outperforms DQN and bootstrapped DQN in terms of exploration efficiency.

1. Introduction

IoT analytics heavily relies on reinforcement learning techniques. Making decisions consecutively is the foundation of reinforcement learning. Simply said, every next input depends on the output of the previous input and the output depends on the state of the current input. Model-based techniques may identify a solution analytically without actually engaging with the environment and can generate the precise result of every state and action interaction in RL. Because the environment is frequently too complicated to create a model, the majority of real-world problems lack models. Second, model-free solutions in RL can only monitor the behavior of the environment through real interaction with it. Machine data are often fragmented and/or include a time aspect. Also, when we accept the data coming

from a certain device, gadgets can act differently depending on the context. Therefore, it is tricky to catch all instances for the data and the data cleaning phase of an algorithm. Continuous sensor data monitoring is also time-consuming and pricey. Reinforcement learning technologies embedded with IoT can help these hazards. These algorithms pick up new information on their own, freeing the programmer to focus on other important responsibilities rather than bothering about the training procedure and data collecting procedure. These developments in computational intelligence methodologies have the potential to greatly improve reliability in terms of latency and processing resources without sacrificing service quality. However, the explosive increase in the use of sensors and actuators for next-generation IoT systems generates enormous quantities of information that are processed all through the cloud, which can materially reduce processing efficiency. Sensor technology is evolving at a rate never witnessed before, driven by advances in materials' science and nanotechnology. As a result, it is becoming more accurate, smaller, and less expensive and capable of detecting things that were not imaginable before. In fact, in a few years, we will see a trillion new sensors deployed yearly since sensing technology is advancing so quickly. Actuator is a different kind of transducer that is used in many IoT systems. Simply said, an actuator works the opposite way from a sensor. It transforms an electrical input into physical movement. Actuators come in a variety of forms, such as electric motors, hydraulic systems, and pneumatic systems. In addition, IoT sensors and actuators can cooperate to provide automation on an industrial scale. Finally, over time, analysis of the data generated by these sensors and actuators might yield insightful business information. Furthermore, the traditional reinforcement learning methods may address exploration issues successfully, but there are certain constraints: the state space of a Markov decision process must be both discrete and constrained; however, in this case, it was not given sufficient attention or focus. Moreover, there are three main problems in the field of reinforcement learning [1]: (1) generalization, (2) how to balance the relationship between exploration and utilization, and (3) credit assignment problem. In recent years, the research hotspot is mainly focused on the generalization problem in large-scale state space, and there is a new research field named deep reinforcement learning [2]. The rise of this field began after the DeepMind team proposed the DQN (deep Q-network) algorithm [3]. The main contribution of the DQN algorithm is to combine deep learning [4] and reinforcement learning [5] and through experience playback (experience replay) and target network mechanism to solve some problems related to generalization [6]. Since then, many improved versions of the DQN algorithm have been proposed [7, 8]. These algorithms mainly focus on the generalization problem, and the main problem to be discussed in this paper is how to balance the relationship between exploration and utilization with the integration of IoT and reinforcement learning. A learning algorithm may be thought of as a stochastic mapping, or a channel in information theory, that takes training data as input and produces a hypothesis as output.

The generalization error is the difference between the output hypothesis' population risk and its empirical risk on the training data. It quantifies the degree to which the learnt hypothesis suffers from overfitting. The usual method of measuring generalization error is based on either specific complexity metrics of the hypothesis space. The reciprocal data between the collection of empirical hazards of the available hypotheses and the algorithm's final output may be utilized to efficiently assess and control bias in data analysis, which is analogous to generalization error in learning issues. The total variation data between a random instance in the dataset and the output hypothesis are used to calculate the generalization error in learning problems; however, the approach is limited to bounded loss functions. Another benefit of relating the generalization error to the input-output mutual information is that, unlike the VC dimension or the uniform consistency, which only depends on the hypothesis space or the learning algorithm, the latter quantity depends on all components of the learning problem, including the dispersion of the dataset, the hypothesis space, the learning algorithm itself, and possibly the loss function. The input-output mutual information may be more closely linked to the generalization error than typical generalization-guaranteeing parameters of interest since the generalization error might substantially depend on the input dataset.

IoT analytics heavily relies on deep learning techniques. Machine data are often sparse and/or contain a temporal component. Even when we believe the data coming from a certain device, gadgets might act differently depending on the environment. Therefore, it is challenging to capture all instances for the data preprocessing/training step of an algorithm. Continuous sensor data monitoring is also time-consuming and costly. Deep learning techniques can reduce these hazards. Deep learning algorithms pick up new information on their own, freeing the developer to focus on other important tasks rather than worrying about the training process. A machine learning method called reinforcement learning involves the agent interacting with its surroundings in an effort to optimize the financial reward. The human brain communicates with the outside world and makes use of that interaction to comprehend and survive in that world. Reinforcement learning compares learning about the environment to learning about the human brain and sensory processing system. It is a process where an agent must investigate every aspect of the system to comprehend it. It is not practical in many situations due to the length of time required to converge and obtain an optimal policy. The problem with traditional RL is its dimensionality. The number of factors that an RL agent must learn increases exponentially as the environment gets more complicated. Over the network, IoT links millions of devices. Because IoT devices are so dynamic, a complicated RL may constantly acquire new data to adapt to many advanced situations. Some IoT ecosystems are so complicated that modeling them is challenging. The effort needed to simulate and solve such a complicated environment is reduced by RL. Think of a challenging IoT case where we need to develop a model to address a challenge.

The hottest supervised learning algorithms in the field of machine learning can only learn relevant patterns from data

collected by humans, but what data need to be collected, or which data are important, is still up to humans to decide. However, IoT-based reinforcement learning can obtain different observation sequences by selecting different actions, and then learning-related patterns according to the obtained observation sequences. That is to say, within the scope of reinforcement learning, an algorithm is needed to decide whether to select the current optimal action (use) or choose an action (exploration) that has the potential to bring long-term benefits. However, algorithms such as DQN use a simple heuristic exploration strategy ε-greedy, which only randomly selects an action with the probability of ε ($0 \leq \varepsilon \leq 1$) to achieve the effect of exploration. However, such an exploration strategy is extremely inefficient. Although the DQN algorithm can reach or even exceed the human level in an Arcade Learning Environment [9], the time complexity of the algorithm is O (2N) when faced with deep exploration problems [4].

A good exploration strategy must consider the information gained brought by the selected action, and the information is mainly measured by the degree of uncertainty of the estimated value, then a good exploration strategy can be obtained only by introducing probabilistic thinking. As early as 1998, Dearden and other scholars proposed Bayesian Q-learning [10] to balance the relationship between exploration and utilization. Since then, other researchers have proposed reinforcement learning algorithms under the Bayesian framework [11, 12]. However, the computational processes of these algorithms are relatively complex and cannot be combined with deep learning techniques. The main problem with the combination of Bayesian reinforcement learning and deep learning is that it is intractable to obtain the posterior distribution of neural network parameters in high-dimensional state space.

The contribution of this paper is to propose a new calculation method that can generate samples of the posterior distribution of parameters, and the IoT-based deep reinforcement learning algorithms using this calculation method are collectively referred to as Bayesian deep reinforcement learning algorithms. Machine intelligence techniques govern computation, regulation, system lag, reliability, consistency, effectiveness, and energy efficiency at multiple elements in IoT ecosystems. These methodologies are in control of information gathering, linking devices to the Internet, data processing, and decision-making without interpersonal interactions. Currently, this calculation method is only applicable to value-based deep reinforcement by the combination of intelligent learning algorithms with policy-based deep reinforcement learning algorithms. Combining the bootstrapped DQN with this calculation method, a new algorithm BBDQN (Bayesian bootstrapped DQN) is obtained, and the exploration efficiency of the algorithm is verified through experiments in two environments, one of which is a lattice world with a deep exploration structure. Another experimental environment is the classic control problem Mountain Car, which does not have a deep exploration structure itself, slightly modify its reward function to make it have a deep exploration structure, and then conduct experiments on the modified Mountain Car.

The experimental results show that BBDQN can solve the deep exploration problem, and its exploration efficiency is better than the DQN algorithm using a random exploration strategy (i.e., ε-greedy) and the bootstrapped DQN algorithm. Through the use of randomized value functions, the straightforward algorithm bootstrapped DQN explores in a computational and statistically efficient way. Bootstrapped DQN engages in temporally extended (or deep) exploration as opposed to dithering tactics such as epsilon-greedy exploration; this can result in tenfold quicker learning. We illustrate these advantages in the expansive arcade learning environment and complicated stochastic MDPs. Bootstrapped DQN significantly reduces learning times and enhances performance in the majority of Atari games. In a traditional bootstrap, you simulate sampling your data from a population by collecting samples with substitutions from your data. This procedure is repeated K times to mimic drawing your sample several times, allowing you to assess the potential variability estimate of your statistic (a function) over various samples drawn from the same population (Xk). Consequently, we are emulating the statistic's "sampling distribution" (variability due to the sampling process). In the specified Bayesian bootstrap, you are estimating the posterior distribution of the estimates of the statistic ϕ (X) and the distribution of your data $X = \{x1, x2, \ldots, xN\}$. It is a nonparametric model in which your data points are assumed to have a categorical distribution (the likelihood).

Even though the DQN has already considerable success, it may still be enhanced with the aid of a variety of different methods, including a more precise approximation approach for the state action-value function and prioritized experience replay. Due to its capacity to balance exploration and exploitation in reinforcement learning, the exploration approach should be the most crucial of these strategies. The exploration strategy determines the course of action to pursue given the current situation and has the potential to affect how the process will ultimately develop in the future. Given its significance, several exploration tactics have been suggested. The strategy employed in the first DQNs and possibly the one that is utilized the most is the epsilon-greedy one. The greedy approach mostly selects the action that is thought to be the best, but it also randomly selects an action with little probability. An activity that has been proven to be a poor option may be repeatedly picked in the succeeding decision-making process since the activities are chosen entirely at random for investigation. To increase the effectiveness of exploration, add some Gaussian noise. For jobs with well-defined incentives, these heuristic strategies are enough.

A significant issue for reinforcement learning is still effective exploration in complicated situations. We suggest the bootstrapped DQN, a straightforward technique that uses randomized value functions to explore in a way that is both computationally and statistically efficient. The bootstrapped DQN engages in temporally extended (or deep) exploration as opposed to dithering tactics such as epsilon-greedy exploration; this can result in tenfold quicker learning. We illustrate these advantages in the expansive Arcade Learning Environment and complicated stochastic

MDPs. The bootstrapped DQN significantly reduces learning times and enhances performance in the majority of Atari games. The bootstrapped DQN builds many posterior estimates using the same original data. The random initialization of the weights for the heads is what causes the variety of approximate solutions. This implies that these heads first try random actions (due to varied random initials) but that some (but not all) of the heads will learn from it due to bootstrapping when some heads discover a favorable state and generalize to it. Other heads will eventually either discover additional excellent states or come to understand the greatest good states discovered by the other heads. Simply employing the bootstrapped DQN as an exploitative approach is quite good in and of itself, even better than the vanilla DQN. This is due to bootstrapped DQN's deep exploration capabilities, since it may employ the best states, it knows while simultaneously planning to try out states it does not know anything about. Even in the films, it is clear that the heads agree on all critical decisions but disagree on other, less significant actions.

2. Related Work

The exploration efficiency of the reinforcement learning algorithm directly affects the sample efficiency of the algorithm, and improving the exploration efficiency can reduce the number of time steps required to train the agent. The Bayes policy is a criterion for measuring the exploration efficiency of reinforcement learning algorithms; however, computation of Bayesian optimal policies is tricky because the computation time grows exponentially with the problem space (states and actions) [13]. There are many studies considering exploration efficiency, and Liu et al. [14] were the authors to propose it, and their work confirmed that polynomial-time reinforcement learning algorithms must use multiperiod exploration. Based on their work, a series of tabular reinforcement learning algorithms has been proposed [15–21]. The exploration methods mentioned in these papers are more efficient than ε-greedy and Boltzmann exploration, but these methods all cannot handle the curse of dimensionality. Deep reinforcement learning algorithms [3, 22, 23] are proposed in recent years to deal with large state spaces tend to use ε-greedy exploration. These algorithms can get good results in Atari arcade games and Go, but they do not have any practical applications; the inefficient exploration efficiency makes these algorithms only trained in simulated environments. By randomly selecting between exploration and exploitation, epsilon-greedy is a straightforward strategy for balancing exploration and exploitation. Epsilon refers to the likelihood of deciding to explore, and the epsilon-greedy, with a low probability of exploring, exploits most of the time. Theoretical computer science, optimization, machine learning, and decision theory all use exponential weighting schemes as basic tools. Exponential weighting systems, often known as Boltzmann, Gibbs, or softmax exploration strategies, are frequently employed in reinforcement learning to balance exploration and exploitation. The chance of selecting an arm in the most popular variant of

Boltzmann exploration is inversely related to an exponential function of the empirical mean of the reward for that arm.

For the exploration of large-scale state space, some researchers proposed a model learning algorithm [24]. The problem with this method is that it can only solve simple model problems. For complex models, the calculation of this method is also difficult. Some researchers have proposed a policy learning algorithm [25], which mainly solves the problem of continuous action space, such as robot control, but when the size of the policy space is exponential, this method cannot guarantee the efficiency of exploration. Another class of approaches encourages exploration by assigning rewards to infrequently visited states based on pseudocounts [26] or density models [27].

How to ensure the exploration efficiency, and generalization of the algorithm has always been a difficult problem in reinforcement learning, and some researchers have thought of using Bayesian thinking to deal with this problem, such as the Bayesian Q-learning algorithm [10], but this algorithm can only deal with state A limited number of problems; there is also the RLSVI (randomized least-square value iteration) algorithm proposed in this literature [28], which is only suitable for linear function approximates and cannot be combined with neural networks, a nonlinear function approximate; while the algorithm proposed in this literature [29] uses the neural network as the feature extractor that takes the input of the last layer of the network as a feature and then uses Bayesian linear regression to calculate the Q-value. RLSVI works similarly to least-squares value iteration (LSVI) and has many of the same principles as other closely related algorithms such as TD, LSTD, and SARSA. The difference between RLSVI and the alternatives is that the algorithm explores by randomly sampling statistically plausible value functions, whereas the alternatives are typically used in conjunction with action-dithering schemes such as Boltzmann or epsilon-greedy exploration, which results in highly inefficient learning. The randomized least-squares value iteration (RLSVI) technique rapidly investigates and generalizes value functions with linearly parameterized values. It is, however, reliant on hand-designed state representation, which necessitates engineering effort for each scenario.

Research in Bayesian neural networks has not been mainstreaming [30, 31]; however, there have been recent signs of a resurgence, with many concerns about quantifying the uncertainty of datasets. Methods [32] have been proposed, and the authors of this paper are inspired by these methods and propose a deep reinforcement learning algorithm that can efficiently compute Bayesian posteriors. Because they provide us with whole distribution across the regression parameters, Bayesian regression methods are known to be particularly powerful. Bayesian linear regression offers a built-in technique for calculating insufficient or unevenly distributed data. The coefficients can have a prior applied to them so that, in the absence of data, the prior can be used instead. In Bayesian linear regression, statistical analysis is carried out under Bayesian interface conditions. To create linear regression, we do not utilize

point estimates but rather probability distributions. An output is produced using a probability distribution rather than more traditional methods of regression. Bayesian linear regression seeks to identify posterior rather than model parameters.

3. Background Knowledge

A collection of potential world states S, a set of models, a set of potential actions A, a real-valued reward function $R(s, a)$, and a policy are all components of a Markov Decision Process (MDP) model. Each possible state for the agent is represented by a set of tokens called a state. A model, often known as a transition model, describes how an action changes a state. Specifically, the transition $T(S, a, S')$ describes a situation in which being in state S and performing an action "a" transports us to state S' (S and S' may be the same). All conceivable actions are included in an action A. The collection of actions that can be executed while in state S is defined as $A(s)$. A reward is a reward function with a real value. The reward for merely existing in the state S is denoted by $R(s)$. The reward for existing in a state S and doing an action "a" is indicated by the expression $R(S, a)$. The reward for being in a state S, doing an action "a," and then ending up in a state S' is denoted by the formula $R(S, a, S')$. The Markov Decision Process has a solution in the form of a policy. A mapping from S to a is a policy. It signifies that when in state S, action "a" should be conducted. Reinforcement learning problems are usually modeled as Markov Decision Processes. The Markov Decision Process used in this paper can be expressed as (S, A, T, R, γ), where S is the state space, A is the action space, and $A(s), s \in S$ is the action set available in state s; T is the state transition probability, $T(S_{t+1}|s_t, a_t)$ represents the probability of transitioning to state S_{t+1} after taking action at state s_t; R is reward functioned, $R_{s,a}$ denotes the mean value of the reward distribution obtained after taking action a in state s, and let r denote the immediate reward obtained after taking action a in state s (it can also be understood as a sample obtained from the reward distribution), and then there is $R_{s,a} = E[r|s_t = s, a_t = a]$; $\gamma \in [0, 1]$ as the discount parameter used to control the weight of future rewards.

The goal of reinforcement learning is to learn a policy that maximizes the discounted cumulative reward obtained by the agent. Let π denote the policy and $\pi(a|s)$ denote the probability of taking action a in state s. Then, introduce the concept of action-value function (Q function), the action-value function $Q_\pi(s, a)$ table under the policy π, and the agent can obtain the expected discounted cumulative reward after taking action states, namely,

$$Q_\pi(s, a) = E_\pi \left[\sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s, a_0 = a \right]. \tag{1}$$

This equation can also be expressed recursively:

$$Q_\pi(s, a) = R_{s,a} + \gamma \sum_{s^t, a^t} \left(s'|s, a \right) \pi \left(a'|s' \right) Q_\pi \left(s', a' \right). \tag{2}$$

We denote the optimal policy by π^* and the optimal action-value function by Q^*. Since the optimal policy is the policy that obtains the maximum discounted cumulative reward, the Q^* function can be written as follows:

$$Q^*(s, a) = R_{s,a} + \sum_{s^t} T\left(s'|s, a \right) \max_{a'} Q^* \left(s', a' \right). \tag{3}$$

Equation (1) is also known as the Bellman optimality equation. For each pair of s, a, there is an equation. By combining all these equations, a system of equations is obtained, which can be solved by linear programming methods. This system of equations yields the optimal action-value function $Q*(s, a)$ for all s, a. The optimal action-value function can also be solved by dynamic programming methods such as policy iteration or value iteration. The optimal strategy can be expressed by the optimal action-value, that is, for all states s, select the action $a = \arg\max a\, Q*(s, a)$ that maximizes the optimal action-value function in this state; if there are multiple actions satisfying with this condition, one of these actions can be randomly selected. It can be seen from equation (1) that this method requires a known environment model—state transition probability T and reward the function R to be calculated. If the environment model is not known, other methods are needed to find the maximum value. Optimal strategies, such as Q-learning algorithms, are used to calculate the optimal action-value function, iteratively, and the update rule is

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[r + \gamma \max Q\left(s', a' \right) - Q(s, a) \right]. \tag{4}$$

Among them, α is the learning rate, and the reward r and the next state, s are both fed back to the agent by the environment. In the case of limited state space and action space, as long as the learning rate sequence satisfies the random approximation condition and all state-action pairs are in the case of continuous updates, the algorithm converges to the optimal value function Q^* [5] with probability 1.

In large-scale state spaces, the traditional reinforcement learning algorithm Q-learning is no longer applicable, but this problem can be solved by combining function approximation techniques. The combination of neural networks and reinforcement learning is a hot research topic in recent years, and the DQN algorithm is a model in this regard. When using a neural network, the action-value function can be expressed as $Q_\theta(s, a)$, where θ is a parameter of the neural network. We compute the target value from feedback from interacting with the environment:

$$y = r + \gamma \max Q_\theta \left(s', a' \right). \tag{5}$$

With the target value and the predicted value, the parameter θ of the neural network can be updated by the stochastic gradient descent algorithm. The update rule is

$$\theta \leftarrow \theta - \alpha \nabla_\theta \left[y - Q_\theta(s, a) \right]^2. \tag{6}$$

If you do this directly, the effect is not very good because the neural network is a supervised learning method, which requires that each training sample is independent of each other, and the samples collected by reinforcement learning are related to each other, and each update of the parameters will affect the target value, causing the target value to be unstable. As mentioned in the DQN algorithm, only two mechanisms are needed to alleviate this problem: experience replay and the target network. The experience replay mechanism is that, each time the agent interacts with the environment, the information (s, a, r, s') of the interaction is stored in the memory pool (replay buffer), and the agent randomly selects from the memory pool for each update. We take a certain number of samples and use these samples to update the parameters. The target network mechanism is that the agent maintains two Q networks, one network is used to select actions and update them in real time and the other network (target network) is used to calculate the target value and will not update in real-time the target value used in each update, and it can be expressed as

$$\gamma = r + \gamma \max Q^{\text{target}}\left(s', a'\right). \tag{7}$$

Among them, Q^{target} represents the target network, and only after the specified time step has passed, the parameters of the two networks will be synchronized once.

4. In-Depth Exploration of the Problem

The chain problem mentioned is a kind of deep exploration problem, and this paper will use this problem to analyze the exploration efficiency of the ε-greedy strategy and Boltzmann exploration. To simplify the calculation, some modifications have been made to the chain problem as discussed in the literature [10]. ε-greedy exploration is to randomly choose one of all available actions with probability ε, and choose the best action for the moment with probability $1 - \varepsilon$. The probability of selecting each action in Boltzmann's exploration is proportional to the estimate of the action-value function, which is calculated as

$$\pi\left(a|s\right) = \frac{e^{Q'(s,a)/t}}{\sum_i e^{Q'(s,a)/t}}. \tag{8}$$

Among them, Q is the estimation of the action-value function, t is the temperature parameter, the larger the value of t is, the closer the strategy is to the random strategy, and the smaller the value of t is, the closer it is to the greedy strategy. A chain problem is used to analyze the exploration efficiency of ε-greedy exploration and Boltzmann exploration. The state transition diagram of this problem is shown in Figure 1. Each arrow in the diagram corresponds to an action and reward. The starting state of the problem is 1, and the termination state is N. Suppose the length of each episode (episode) $H = N - 1$, and no matter in which state, the choice of action a will have, a probability of 0.2 failure (failure means that the selected action and the actual action are inconsistent) and action b is deterministic. The optimal strategy for this problem is to always choose action a, and the
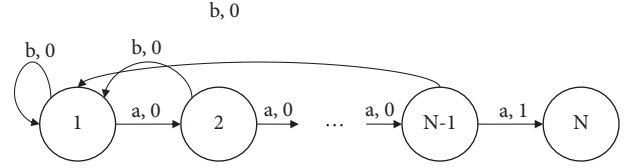
Figure 1: Chain problem.

average reward that the optimal strategy can get in each cycle is the probability $(1 - 0.2)^{N-1}$ that the strategy successfully reaches N. Whether it is ε-greedy or Boltzmann exploration, since there is no reward before reaching the target state N (i.e., the reward is 0), the estimated values of action a and action b are equal in all states, and the action selection is completely random; in this case, the probability of reaching state N through exploration is $(1 - 0.2)^{N-1} \times 2^{-(N-1)}$, just take the reciprocal of this probability to get the average required amount of reaching state N through exploration. The number of cycles, with l representing the required number of cycles, the following relationship is established:

$$\text{the } E[l] = \frac{1}{(1 - 0.2)^{n-1}} 2^{N-1} > 2^{N-1}. \tag{9}$$

From this inequality, it can be determined that a reinforcement learning algorithm using an ε-greedy strategy or Boltzmann exploration has a time complexity of $O(2^N)$ when faced with a deep exploration problem such as the chain problem, where N represents the size of the problem.

The exploration efficiency of epsilon-greedy exploration and Boltzmann exploration is examined using a chain problem. Figure 1 depicts the state transition diagram for this problem. Each arrow in the graphics represents an activity and its associated reward. The problem's initial state is 1 and its final state is N. The described chain issue is a type of deep exploration problem, and this study will utilize it to compare the exploration efficiency of the epsilon-greedy method with Boltzmann exploration. Epsilon-greedy exploration is the process of randomly selecting one of all potential actions with probability, and then selecting the optimal action for the moment with probability. In Boltzmann's exploration, the likelihood of picking each action is proportional to the estimated action-value function.

5. Bayesian Deep Reinforcement Learning

The benefits of exploration can be estimated by the value of information. The so-called value of information refers to the degree to which the information obtained through exploration leads to the improvement of the quality of future decision-making, and how quantifying the information obtained through exploration is the key. According to information theory, the amount of information brought by the exploration can be calculated by the uncertainty of the estimated value of the action selected by the exploration. The quantitative treatment of the ideas, parameters, and regulations regulating message transmission through communication networks is known as information theory. It was established by Claude Shannon in the middle of the 20th

century and has since developed into a vibrant branch at the intersection of communication theory and mathematics, promoting the advancement of other scientific disciplines. Some people consider information theory to be a subset of probability theory since the methods utilized in it are probabilistic in nature. Whether an analog or digital communication technology is used, information is its source. Information theory is a mathematical method for studying information coding, as well as quantification, storage, and communication. In [10], a Bayesian method is proposed to maintain this information; however, in this paper, the number of state-action pairs of the problem discussed is limited, and all the obtained state-action pairs can be recorded by recording all the variance and then calculate the amount of information. When the number of state-action pairs is too large, reinforcement learning will consider combining generalization techniques to solve this problem. If a linear function approximator is used, it can be estimated according to the Bayesian linear regression mentioned in [33] value variance. However, in recent years, research related to deep learning has shown that the strong generalization ability of neural networks is much stronger than that of linear function approximators. Therefore, the combination of deep learning and reinforcement learning technology has become a major trend, and a new field of research is called deep reinforcement learning. Due to the complex structure of the neural network and the huge number of parameters, the variance of the estimate cannot be calculated by the calculation method mentioned in the literature [33], so this paper proposes a new calculation method. This method obtains a sample of the posterior distribution of the neural network parameters by calculation at each time step. Since the estimated variance of different actions is different, the action with high variance has a larger amount of information, and its sampling value may also be larger. Therefore, the probability of being selected is also higher. For example, there are two actions whose Q-value obeys a Gaussian distribution. Since action 1 is selected many times, its variance is smaller than that of action 2. Suppose that the variance of action 1 is 1, and the variance of action 2 is 10. In addition, the variance of action 1 is 1. Action 1 has a mean historical return of 3 and action 2 has a mean historical return of 1, so action 1 follows a Gaussian distribution with mean 3 and variance 1, while action 2 follows a Gaussian distribution with mean 1 and variance 10, resulting in action 2. Due to the large variance, the probability that its sampled value is greater than that of action 1 is also greater. The following sections describe in detail how to

apply Bayesian methods in deep reinforcement learning algorithms.

The input to a neural network can be a one-dimensional array, a two-dimensional matrix, or even a three-dimensional image. For the sake of clarity, $x \in \mathfrak{R}^d$ is used to represent the state feature, that is, the input value of the neural network, plus the target value y to form the training set, which is expressed as $D = \{(x_i y_i)\}_{i=1}^n$. The posterior distribution of the linear regression model $f_\theta(x_i) = \theta^T x_i$ parameters is calculated and later extended to the case of neural networks. The model parameters are expressed as $s \in \mathfrak{R}^d$, and the prior distribution of the parameters θ is assumed to be $N(\theta^-, \lambda I)$, and the observed target value has a certain noise, that is, $y_i = \theta^T x_i + w_i$, where $w_i \propto N(0, \sigma^2)$. For noise, the noise of each sample is independent of each other. According to Bayes' theorem, the posterior distribution of the parameter θ can be expressed as

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}. \tag{10}$$

In Bayesian theory, $p(D|\theta)$ is called the likelihood, $p(\theta)$ is called the prior, and $p(D)$ is called the evidence. According to the derivation of the literature [33], the posterior distribution of the parameter θ obeys the multivariate Gaussian distribution with the mean (12) and the covariance (13):

$$E[\theta|D] = \left(\frac{1}{\sigma^2}X^T X + \frac{1}{\lambda}I\right)^{-1}, \tag{11}$$

$$\text{cov}[\theta|D] = \left(\frac{1}{\sigma^2}X^T X + \frac{1}{\lambda}I\right)^{-1}. \tag{12}$$

Among them, $X \in \mathfrak{R}^{n \times d}$ is the matrix obtained by splicing n sample inputs $x_i i$ and $y \in \mathfrak{R}^n$ is the vector obtained by splicing n target values y_i. This posterior distribution can be inferred by methods such as Markov Chain Monte Carlo, but to extend to nonlinear model artificial neural networks, the posterior distribution is expressed in another form:

$$\theta' \leftarrow \left(\frac{1}{\sigma^2}X^T X + \frac{1}{\lambda}I\right)^{-1}\left(\frac{1}{\sigma^2}X^T(y+\delta) + \frac{1}{\lambda}\theta'\right). \tag{13}$$

Among them, $\delta \in \mathfrak{R}^n$ is a random vector, each component δ_i comes from the Gaussian distribution $(0, \sigma^2)$, and the sampling of each component is independent of each other, and θ comes from the parameter prior distribution $N(\theta', \lambda I)$. It can be shown that θ' and θ are equivalent because

$$E\left[\theta'|D\right] = \left(\frac{1}{\sigma^2}X^T X + \frac{1}{\lambda}I\right)^{-1} \cdot \left(\frac{1}{\sigma^2}X^T(y + E[\delta|D]) + \frac{1}{\lambda}E\left[\theta'|D\right]\right) = E[\theta|D]. \tag{14}$$

That is, the means of θ' and the mean of θ are equal, and there are

$$\text{cov}[\theta|D] = \left(\frac{1}{\sigma^2}X^TX + \frac{1}{\lambda}I\right)^{-1}.\left(\frac{1}{\sigma^4}X^TE[\delta\delta^T|D]X + \frac{1}{\lambda^2}E\left[\theta'\theta'^T|D\right]\right). \tag{15}$$

$((1/\sigma^2)X^TX + (1/\lambda)I)^{-1} = \text{cov}[\theta|D]$ and $E[\delta\delta^T|D] = \sigma^2 I$, $E[\theta'\theta'^T|D] = \lambda I$, that is, the covariance of θ' is equal to the covariance of θ. With θ', the posterior parameters can be directly calculated, which is the key to combine Bayesian methods and neural networks. Calculating θ' requires adding noise to the target value, which can be obtained by open delta number generation noise sample open δ_i for each sample (x_i, y_i). In addition, the reinforcement learning algorithm will continuously collect new data and modify its model parameters over time. Therefore, formula (15) is rewritten into an iterative calculation method to adapt it to the form of online learning. The update rules are as follows:

$$\theta' \leftarrow \theta' - \nabla_\theta L(f_\theta + f_\theta; D_{\text{noise}}). \tag{16}$$

Among them, $\nabla_\theta L(f_\theta + f_\theta; D_{\text{noise}})$ is the loss function and $f_{\theta'} + f_\theta$ indicates that the output of the prediction function is obtained by adding the outputs of two neural networks, one of which has the posterior parameter θ' and the other network has the prior. The parameter θ'; D_{noise} represents the dataset with added noise, i.e., $D_{\text{noise}} = \{(x_i y_i + \delta_i)\}_{i=1}^n$. Equation (16) is to update the posterior parameter θ', and the prior parameter θ will not change after it is determined in the initialization stage. We combine formula (16) for calculating the posterior parameters and the bootstrapped DQN method, since the DQN method adopts the experience replay mechanism, and the way of sampling samples from the memory pool in the experience replay mechanism is equivalent to inject Gaussian noise into the data [34], so there is no need to explicitly add noise; the dataset is represented by $D_{\text{sample}} = \{(x_i y_i + \delta_i)\}_{i=1}^n$, where n represents the number of samples drawn from the memory pool at each update, also known as the batch size. We denote the target network as $f_{\theta'}$, and then the loss function can be written as

$$\text{opensL}(f_{\theta'} + f_{\theta'}; D_{\text{sample}}) := \sum_{t \in D_{\text{sample}}} (y_i - (f_{\theta'} + f_{\theta'})(x_i)). \tag{17}$$

The method that combines the Bayesian method and bootstrapped DQN is called the BBDQN algorithm, see Algorithm1 for an overall description of the algorithm.

5.1. BBDQN Algorithm.

The Bayesian bootstrap, like other bootstrapping approaches, can enhance probabilistic forecasts by employing preprocessed data with replacement, which allows for distinguishing the output predictions. The Bayesian bootstrap in this research works analytically to obtain the posterior distributions of the QR model parameters, distinguishing itself from the classic bootstrap, which depends on random choices from the available input data. Because the observations are treated as fixed in Bayesian inference, this new perspective appears to be compatible with the Bayesian method. Indeed, the determination of the mean using the Bayesian bootstrap varies only in the weight distribution. However, from a conceptual viewpoint, the Bayesian bootstrap differs significantly from the frequentist form.

6. Experimental Analysis

6.1. Lattice World.

First, test the exploration efficiency of the BBDQN algorithm in the grid world shown in Figure 2. The white grid represents the scale of the grid world, and the gray grid is the terminal state. Due to space constraints, the grid world in Figure 2 has a scale of only 4×4, while the scale of the lattice world used in the experiment is 20×20, but this does not prevent the use of the small-scale lattice world of Figure 2 to describe its dynamic model. In Figure 2, the state S is the initial state, and the agent has always maintained a speed of moving to the right $+1$. The available actions in each state are up and down. If you choose up, then the next step will reach the upper right state of the current state, and if down is selected, the next step will reach the lower right state of the current state. If the agent is at the bottom of the grid world, the downward action can be understood as walking against the wall. At this time, the next state will be to the right of the current state. If you select the action up in the upper right corner of the white grid, you can reach the top of the gray grid. To get a reward of $+1$, there are no rewards in other states, so you must always choose the action up to get the reward, but the action up has a cost, and the cost is related to the size of the grid world. Assuming that the size of the grid world is $N \times N$, each selection action up will bring a reward of $-0.01/N$, while choosing action down has no cost, the reward is 0. In fact, this problem is a two-dimensional extended version of the chain problem mentioned in Paper 4. The input can be represented as a one-hot matrix $x_i \in \{0, 1\}^{N \times N}$, the position of the agent in the matrix is 1, and the other positions are all 0s. In this experiment, compared with the DQN using ε-greedy strategy and bootstrapped DQN, the hyperparameters used by BBDQN algorithm are shown in Table 1, where $0 \in \mathfrak{R}^d$ represents a vector with all 0 components.

In the grid world, simple is the norm. The left-bottom corner of your agent or robot serves as the "start" point, and it terminates at either $+1$ or -1, which corresponds to the associated reward. The agent has four alternative movements at each step, including up, down, left, and

(1) Input: di scount parameter γ, batch size n, number of learners K, parameter first
(2) Test mean $\bar{\theta}$, parameter prior variance λ, sampling interval T sample, target network more
(3) The new interval T target.
(4) Output: network parameters $\bar{\theta}$.
(5) Initialize the parameters $\bar{\theta}$ of the K learners.
(6) Initialize the memory pool and set count = 0
(7) The prior parameters $\hat{\theta}$ of K learners are obtained by sampling
(8) for episode = 1 to inf do
(9) Get the initial state $x1$ from the environment
(10) for $t = 1$ to T do
(11) if count mod T sample = 0 then
(12) Sample $k \sim \text{Unif}(\{0,1,\ldots,K\})$
(13) end if
(14) Choose the optimal action at accordi ng to the learner k
(15) Execute the action at to get the rt and $xt + 1$ of the environmental feedb ack
(16) store $(xt, at, rt, xt + 1)$ into memory pool
(17) Draw n da ta samples from the memory pool to get Dsample
(18) The loss function is calculated by formula (6):
(19) $\bar{\theta} \leftarrow \bar{\theta} - \alpha \nabla \theta L(f\bar{\theta} + f\hat{\theta}; \text{Dsample})$
(20) if count mod T target = 0 then
(21) Upda te target network
(22) end if
(23) count = count + 1
(24) end for
(25) end for

ALGORITHM 1: Bootstrapped DQN (BBDQN) algorithm.
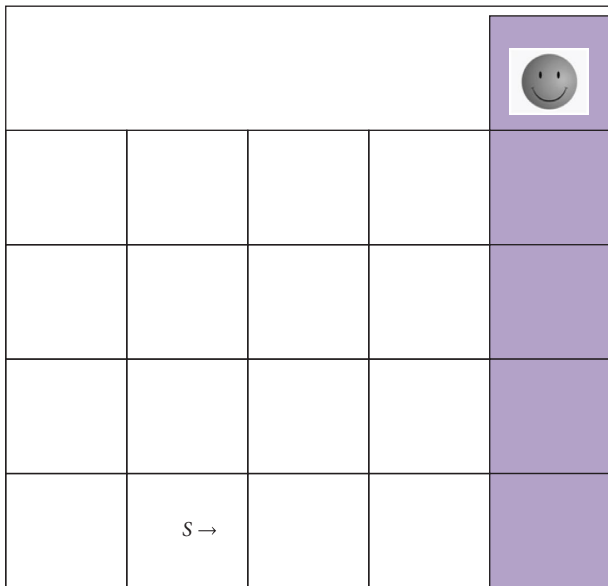
FIGURE 2: Grid world with deep exploration structure.

TABLE 1: Hyperparameters.

Hyperparameters	Value
Discount parameters γ	0.99
Batch size n	128
Memory pool capacity	100
Number of learners K	10
Parameter prior mean $\bar{\theta}$	0
Parameter prior variance λ	10
Sampling interval T^{sample}	20
Target network update interval T^{target}	20

right, while the black block is a wall that your agent cannot pass through. Our first solution assumes that each action is deterministic, i.e., the agent will move in the direction it desires to mov in order to simplify things. The wall, on the other hand, will stay in place if the agents run into it. This is where artificial intelligence comes into play, because our bot should be able to learn from the procedure and think like a person. Value iteration is the key to the magic. Once our agent discovers a path to reward +1, should it stick to it and always take that line (exploitation), or should it give other paths a shot (exploration) and expect a shorter path? In reality, we will balance exploitation and exploration in order to keep our agent from being stuck in the local optimum. Our agent will pick an action depending on the exploration rate.

The experimental results are shown in Figure 3, where the algorithm performance is measured with regret. Regret refers to the difference between the maximum cumulative reward and the actual cumulative reward. Because DQN has never found a rewarded area in the grid world, its cumulative reward is less than 0, and its regret is constantly growing. As can be seen from Figure 3, the BBDQN can learn the optimal policy in less than 2000 epochs, while the DQN does not learn a good policy after 10000 epochs. In fact, after 1 million cycles of experiments, the DQN still did not converge.

FIGURE 3: Algorithm performance over epochs for rewards gain.

TABLE 2: The number of episodes required to find rewards.

N	l
5	560
10	792
15	1663
20	1828
25	2657
30	7697

Among them, the parameter retains two decimal places. Therefore, when the BBDQN algorithm faces deep exploration problems, the time complexity of the algorithm is at the polynomial level, specifically $O(N^3)$, which is better than the random exploration strategy ε-greedy time complexity of $O(2^N)$.

6.2. Further Analysis on Lattice World. In Figure 3, there is no obvious difference between the DQN algorithm and the bootstrapped DQN algorithm, but this does not mean that the exploration efficiency of the two algorithms is the same. In fact, the bootstrapped DQN algorithm is more efficient than the DQN. To prove this conclusion, a series of experiments were carried out.

In the previous section, it was mentioned that the size of the lattice world is represented by $N \times N$, and the number of learners is represented by K. The performance of each algorithm is compared when K is fixed at 10 and N is 10, 20, and 30. In addition, the performance of each algorithm is compared when N is fixed at 20 and K is 10, 20, and 30. When the lattice size of the world becomes smaller, the superiority of the bootstrapped DQN over the DQN is reflected. As shown in Figure 4, when K is 10 and N is 10, the regret of the DQN algorithm keeps growing, while the bootstrapped DQN and BBDQN find the optimal policy within 1000 cycles. In addition, when the number of learners increases, the superiority of the bootstrapped DQN over the DQN can also be reflected. As shown in Figure 5, when K is 30 and N is 20, the bootstrapped DQN algorithm finds the optimal policy at around 2000 cycles.

The results of all experiments are shown in Table 3, where the bootstrapped DQN is denoted as the BDQN due to space constraints. In addition, the experimental results in Section 6.1 are the results shown in the table with $K = 10$ and $N = 20$. In Table 3, the algorithm can learn the optimal policy in 10,000 epochs.

Regrettably shown in bold, it can be seen that the DQN fails to find the optimal policy within 10,000 epochs in all experimental settings, while the bootstrapped DQN only works when $K = 10$, $N = 10$ (the size of the lattice world decreases), or $K = 30$ and $N = 20$ (the number of learners increases); the optimal policy can be learned in 10,000 epochs, while the BBDQN can learn the optimal policy in 10,000 epochs under all settings. It can also be seen from Table 3 that when the size of the lattice world is small ($N = 10$), the performance of the BBDQN algorithm and the bootstrapped DQN algorithm is not much different because the BBDQN adds a random initialization first.

The time complexity of the DQN algorithm using the ε-greedy strategy is $O(2N)$ when faced with deep exploration problems. This paper also tests the performance of the bootstrapped DQN algorithm. The action selection of this algorithm is not realized by ε-greedy, and the algorithm improves its exploration efficiency by adding ensemble technology on the basis of the DQN [10], but it cannot be used in solving this deep exploration problem in the short term. The regret of the bootstrapped DQN and DQN over 10 000 cycles has only a small difference, which is not visible in Figure 3 and must be zoomed in to see the difference. Specifically, the DQN has accumulated regrets of 9 over 10000 cycles.

Figure 3 is the performance curve of the algorithm when the grid world size is 20×20. It is possible to assume that the size of the lattice world is $N \times N$ and analyze the relationship between algorithm exploration efficiency and N. As in Paper 4, the number of cycles required for the algorithm to discover the reward for the first time is used as a measure of exploration efficiency, and 6 data samples are obtained through experiments, which are shown in Table 2, where l represents the first discovery of the reward by the algorithm. In order to express the relationship between the two, polynomial regression is used to fit these 6 points, which is expressed as follows:

$$l = b + w_1 N + w_2 N^2 + \ldots + w_m N^m. \tag{18}$$

Among them, $w_1, w_2, \ldots, w_m m$ are the parameters of the m degree polynomial, which can be obtained by the least-square method. We tested multiple m values and found that when m is 3, the mean square error is the smallest, that is, the relationship between the two can be approximated by a 3rd degree polynomial as follows:

$$l \approx -2.76 + 9.30N - 6.84N^2 + 1.63N^3. \tag{19}$$

Figure 4: Algorithm performance over epochs for rewards gain ($N = 10$ and $K = 10$).
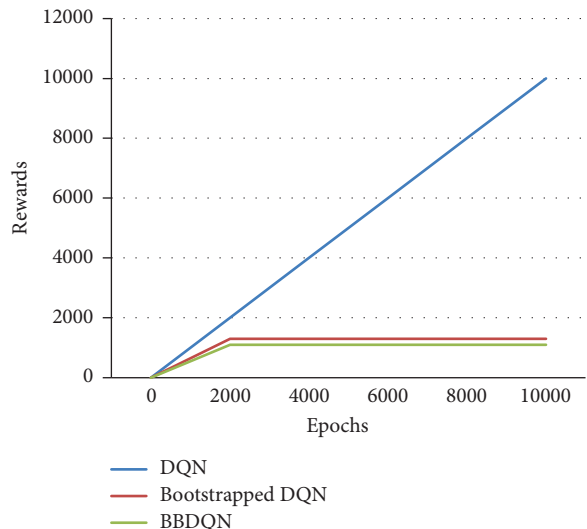
Figure 5: Algorithm performance over epochs for rewards gain ($N = 20$ and $K = 20$).

Table 3: Cumulative regret of each algorithm in 10,000 episodes.

Variable settings		DQN	BDQN	BBDQN
	$N = 10$	9914.88	599.65	649.29
$K = 10$	$N = 20$	9922.02	9914.60	1590.03
	$N = 30$	9912.36	9910.96	5203.65
	$K = 10$	9922.02	9914.02	1590.03
$N = 20$	$K = 20$	9922.02	9912.13	1901.18
	$K = 30$	9922.02	1857.49	1784.71

The performance of the BBDQN is even slightly lower than the bootstrapped DQN, but when the scale of the grid world increases, the advantages of the BBDQN algorithm proposed in this paper become more obvious, which is also the embodiment of the BBDQN more suitable for solving deep exploration problems; in addition, BBDQN algorithm performance of the method does not increase with the number of learners, which means that the space requirement of the BBDQN is lower than that of the bootstrapped DQN because the more the number of learners, the more space is required to store each learned parameter, and these algorithms all use neural networks as function approximators, and the parameters of each network are in millions, and the BBDQN does not need to increase the number of learners like the bootstrapped DQN to improve algorithm performance (as shown in Table 3), so the space requirement of the BBDQN is lower than the bootstrapped DQN.

7. Conclusion

A variety of learning and optimization methods form the foundation of computational intelligence. Therefore, incorporating cutting-edge learning methods to balance the relationship between exploration and exploitation is an inspiring area, particularly when merging it with IOT. As a result, IoT collected the scattered data and screened the information out of it using computational algorithms, and reinforcement learning performed the deep exploration with the ultimate emphasis. While disregarding the issue of identifying and taking advantage of the dilemma, the reinforcement learning approaches developed in recent years have mostly concentrated on integrating deep learning technology to enhance the algorithm's generalization capabilities. In order to increase the effectiveness of exploration, a deep reinforcement algorithm based on computational intelligence is proposed in this study, using intelligent sensors and the Bayesian approach. In addition, the technique for computing the posterior distribution of parameters in Bayesian linear regression is expanded to nonlinear models such as artificial neural networks; also, this calculation method is combined with the bootstrapped DQN to obtain the BBDQN algorithm. The result of the experiments in two environments proves that the exploration efficiency of the BBDQN algorithm is better than the DQN algorithm using the ε-greedy strategy and the bootstrapped DQN algorithm.

The direction of further research is when the method of calculating the posterior distribution of the Q function parameter in this paper is used to calculate the posterior distribution of the policy function π_θ parameter in the policy gradient method, the policy gradient method can effectively improve the face of the policy gradient method and exploration efficiency when exploring problems in depth.

Data Availability

The data used to support the findings of this study are available from the corresponding author on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. J. Liston and K. R. Dandekar, "Entropy based exploration in cognitive radio networks using deep reinforcement learning for dynamic spectrum access," in *Proceeding of the*

2021 IEEE 21st Annual Wireless and Microwave Technology Conference (WAMICON), pp. 1–5, Sand Key, FL, USA, April 2021.

[2] N. Dilokthanakul, C. Kaplanis, N. Pawlowski, and M. Shanahan, "Feature control as intrinsic motivation for hierarchical reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3409–3418, 2019.

[3] A. Rkhami, Y. Hadjadj-Aoul, and A. Outtagarts, "Learn to improve: a novel deep reinforcement learning approach for beyond 5G network slicing," in *Proceeding of the 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, Las Vegas, NA, USA, June 2021.

[4] Y. Yao, L. Xiao, Z. An, W. Zhang, and D. Luo, "Sample Efficient Reinforcement Learning via Model-Ensemble Exploration and Exploitation," in *Proceeding of the 2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4202–4208, Xi'an, China, June 2021.

[5] H. Huang, C. Zeng, Y. Zhao et al., "Scalable orchestration of service function chains in NFV-enabled networks: a federated reinforcement learning approach," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2558–2571, 2021.

[6] X. Qiu, W. Zhang, W. Chen, and Z. Zheng, "Distributed and collective deep reinforcement learning for computation offloading: a practical perspective," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 5, pp. 1085–1101, 2021.

[7] W. Liu, Y. He, J. Zhang, and J. Qiao, "Deep reinforcement learning-based MEC offloading and resource allocation in uplink NOMA heterogeneous network," in *Proceeding of the 2021 Computing, Communications and IoT Applications (ComComAp)*, pp. 144–149, Shenzhen, China, May 2021.

[8] B. Li and Y. Wu, "Path planning for UAV ground target tracking via deep reinforcement learning," *IEEE Access*, vol. 8, pp. 29064–29074, 2020.

[9] F. Soleymani and E. Paquet, "Financial portfolio optimization with online deep reinforcement learning and restricted stacked autoencoder—DeepBreath," *Expert Systems with Applications*, vol. 156, Article ID 113456, 2020.

[10] J. Yang and G. Peng, "DDPG with meta-learning-based experience replay separation for robot trajectory planning," in *Proceeding of the 2021 7th International Conference on Control, Automation and Robotics (ICCAR)*, pp. 46–51, Singapore, July 2021.

[11] R. Ceren, K. He, P. Doshi, and B. Banerjee, "PALO bounds for reinforcement learning in partially observable stochastic games," *Neurocomputing*, vol. 420, pp. 36–56, 2021.

[12] M. Abdar, F. Pourpanah, S. Hussain et al., "A review of uncertainty quantification in deep learning: techniques, applications and challenges," *Information Fusion*, vol. 76, pp. 243–297, 2021.

[13] Y. Bao, Y. Peng, and C. Wu, "Deep learning-based job placement in distributed machine learning clusters," *IEEE INFOCOM 2019 IEEE Conference on Computer Communications*, vol. 18, pp. 505–513, 2019.

[14] W. Liu, L. Peng, J. Cao et al., "Ensemble bootstrapped deep deterministic policy gradient for vision-based robotic grasping," *IEEE Access*, vol. 9, pp. 19916–19925, 2021.

[15] A. Menezes, P. Vicente, A. Bernardino, and R. Ventura, "From rocks to walls: a model-free reinforcement learning approach to dry stacking with irregular rocks," in *Proceeding of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 2057–2065, Virtual Conference, June 2021.

[16] K. Yang, R. Cao, Y. Zhou, J. Zhang, E. Shao, and G. Tan, "Deep reinforcement agent for failure-aware job scheduling in high-performance computing," in *Proceeding of the 2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 442–449, Beijing, China, May 2021.

[17] A. Asghari, M. K. Sohrabi, and F. Yaghmaee, "A cloud resource management framework for multiple online scientific workflows using cooperative reinforcement learning agents," *Computer Networks*, vol. 179, Article ID 107340, 2020.

[18] A. Mehbodniya, J. L. Webber, M. Shabaz, H. Mohafez, and K. Yadav, "Machine learning technique to detect sybil attack on IoT based sensor network," in *IETE Journal of Research*, Informa UK Limited, Colchester, England, 2021.

[19] Z. Cao, K. Wong, and C.-T. Lin, "Weak human preference supervision for deep reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5369–5378, 2021.

[20] J. Zhou, J. Luo, J. Wang, and L. Deng, "Cache pollution prevention mechanism based on deep reinforcement learning in NDN," *Journal of Communications and Information Networks*, vol. 6, no. 1, pp. 91–100, 2021.

[21] S. Dong, Y. Xia, and T. Peng, "Network abnormal traffic detection model based on semi-supervised deep reinforcement learning," *IEEE Transactions on Network and Service Management*, vol. 18, no. 4, pp. 4197–4212, 2021.

[22] C. Huang, H. Zhang, L. Wang, X. Luo, and Y. Song, "Mixed deep reinforcement learning considering discrete-continuous hybrid action space for smart home energy management," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 743–754, 2022.

[23] M. Haliem, G. Mani, V. Aggarwal, and B. Bhargava, "A distributed model-free ride-sharing approach for joint matching, pricing, and dispatching using deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 12, pp. 7931–7942, 2021.

[24] Y. Tao, J. Qiu, and S. Lai, "A hybrid cloud and edge control strategy for demand responses using deep reinforcement learning and transfer learning," *IEEE Transactions on Cloud Computing*, vol. 10, no. 1, pp. 56–71, 2022.

[25] K. Zhu and T. Zhang, "Deep reinforcement learning based mobile robot navigation: a review," *Tsinghua Science and Technology*, vol. 26, no. 5, pp. 674–691, 2021.

[26] Y. Zhao, Y. Ma, and S. Hu, "USV formation and path-following control via deep reinforcement learning with random braking," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5468–5478, 2021.

[27] S. K. Prabhakar and S.-W. Lee, "SASDL and RBATQ: sparse autoencoder with swarm based deep learning and reinforcement based Q-learning for EEG classification," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 3, pp. 58–68, 2022.

[28] J. Sharma, P.-A. Andersen, O.-C. Granmo, and M. Goodwin, "Deep Q-learning with Q-matrix transfer learning for novel fire evacuation environment," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 12, pp. 7363–7381, 2021.

[29] J. Dezert, A. Tchamova, and D. Han, "Total belief theorem and generalized Bayes' theorem," in *Proceeding of the 2018 21st International Conference on Information Fusion (FUSION)*, pp. 1040–1047, Cambridge, United Kingdom, July 2018.

[30] F. Ikorasaki and M. B. Akbar, "Detecting corn plant disease with expert system using Bayes theorem method," in *Proceeding of the 2018 6th International Conference on Cyber and*

IT Service Management (CITSM), pp. 1–3, Parapat, Indonesia, November 2018.

[31] S. Tharewal, M. W. Ashfaque, S. S. Banu, P. Uma, S. M. Hassen, and M. Shabaz, "Intrusion detection system for industrial internet of things based on deep reinforcement learning," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 9023719, 8 pages, 2022.

[32] D. Elavarasan and P. M. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, vol. 8, pp. 86886–86901, 2020.

[33] Z. Wen and B. Van Roy, "Efficient exploration and value function generalization in deterministic systems," *Advances in Neural Information Processing Systems*, vol. 1, pp. 3021–3029, 2013.

[34] I. Osband and B. Van Roy, *Bootstrapped thompson sampling and deep exploration*, https://arxiv.org/abs/1507.00300, 2015.

Retraction

Retracted: The Construction of Big Data Computational Intelligence System for E-Government in Cloud Computing Environment and Its Development Impact

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Ma, F. Hao, Y. Lin, and Y. Liang, "The Construction of Big Data Computational Intelligence System for E-Government in Cloud Computing Environment and Its Development Impact," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7295060, 9 pages, 2022.

Retraction

Retracted: A Study of Feature Construction Based on Least Squares and RBF Neural Networks in Sports Training Behaviour Prediction

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Qiu, C. Su, X. Liu, and D. Yu, "A Study of Feature Construction Based on Least Squares and RBF Neural Networks in Sports Training Behaviour Prediction," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5034081, 7 pages, 2022.

Retraction

Retracted: Internet False News Information Feature Extraction and Screening Based on 5G Internet of Things Combined with Passive RFID

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] D. Hu and Q. Xia, "Internet False News Information Feature Extraction and Screening Based on 5G Internet of Things Combined with Passive RFID," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9696472, 11 pages, 2021.

Retraction

Retracted: Financial Market Sentiment Prediction Technology and Application Based on Deep Learning Model

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Guo, "Financial Market Sentiment Prediction Technology and Application Based on Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1988396, 10 pages, 2022.

Retraction

Retracted: Recognition of Continuous Music Segments Based on the Phase Space Reconstruction Method

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Huang and Y. Hu, "Recognition of Continuous Music Segments Based on the Phase Space Reconstruction Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4099505, 11 pages, 2022.

Retraction

Retracted: Application of Improved DEA Algorithm in Public Management Problem Classification

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Li, "Application of Improved DEA Algorithm in Public Management Problem Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2568162, 13 pages, 2022.

Retraction

Retracted: Prevention of Business Risks of Internet Information Security Platforms Based on Blockchain Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] B. Yang, "Prevention of Business Risks of Internet Information Security Platforms Based on Blockchain Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7671810, 10 pages, 2022.

Retraction

Retracted: Teaching Mode Based on Educational Big Data Mining and Digital Twins

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Zhou and X. Wu, "Teaching Mode Based on Educational Big Data Mining and Digital Twins," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9071944, 13 pages, 2022.

Retraction

Retracted: Optimal Design of International Trade Logistics Based on Internet of Things Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Wang, "Optimal Design of International Trade Logistics Based on Internet of Things Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8781095, 11 pages, 2022.

Retraction

Retracted: On the Design of Secured and Reliable Dynamic Access Control Scheme of Patient E-Healthcare Records in Cloud Environment

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] K. Zala, H. K. Thakkar, R. Jadeja et al., "On the Design of Secured and Reliable Dynamic Access Control Scheme of Patient E-Healthcare Records in Cloud Environment," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3804553, 19 pages, 2022.

Hindawi

Retraction

Retracted: DE Oxidation-Fused Industrial Wastewater Purification Fuzzy Control and Simulation

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Ma, "DE Oxidation-Fused Industrial Wastewater Purification Fuzzy Control and Simulation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2843055, 8 pages, 2022.

Retraction

Retracted: A Corpus-Driven Study on Three Elements of Third Personal Pronoun Endophora in English Abstracts of Chinese and Foreign Theses

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Luo, "A Corpus-Driven Study on Three Elements of Third Personal Pronoun Endophora in English Abstracts of Chinese and Foreign Theses," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5981285, 7 pages, 2022.

Retraction

Retracted: Promoting Regional Economic Transformation Forecast Based on Intelligent Computing Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Cai, "Promoting Regional Economic Transformation Forecast Based on Intelligent Computing Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1835376, 12 pages, 2022.

Retraction

Retracted: Design of a Regional Economic Forecasting Model Using Optimal Nonlinear Support Vector Machines

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] T. Zhang, "Design of a Regional Economic Forecasting Model Using Optimal Nonlinear Support Vector Machines," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2900434, 10 pages, 2022.

Retraction

Retracted: Big Data Value Calculation Method Based on Particle Swarm Optimization Algorithm

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] W. Ma and X. Hou, "Big Data Value Calculation Method Based on Particle Swarm Optimization Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5356164, 8 pages, 2022.

Retraction

Retracted: Research and Development of User Clustering-Based Content Similarity Algorithms in Dance-Assisted Choreography Techniques

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Wu and M. Liu, "Research and Development of User Clustering-Based Content Similarity Algorithms in Dance-Assisted Choreography Techniques," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1364835, 10 pages, 2022.

Retraction

Retracted: Ageing Design of Urban Park Landscape Based on Computer Virtual Simulation Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Fan, "Ageing Design of Urban Park Landscape Based on Computer Virtual Simulation Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3150371, 9 pages, 2022.

Retraction

Retracted: English Speech Recognition System Model Based on Computer-Aided Function and Neural Network Algorithm

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Zhang, "English Speech Recognition System Model Based on Computer-Aided Function and Neural Network Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7846877, 11 pages, 2022.

Hindawi

Research Article

Advances in Computational Intelligence Techniques-Based Multi-Intersection Querying Theory for Efficient QoS in the Next Generation Internet of Things

Ashish Kumar ,[1] Kannan K ,[2] Mamta Dahiya ,[3] Virendra Singh Kushwah ,[4] Ayesha Siddiqa ,[5] Kiranjeet Kaur ,[6] and Saima Ahmed Rahin [7]

[1]*Department of CSE, Manipal University Jaipur, Jaipur, Rajasthan, India*
[2]*Department of CSE, Sree Vidyanikethan Engineering College, Tirupati, Andhra Pradesh, India*
[3]*Department of Computer Science and Engineering, SGT University, Gurugram, India*
[4]*School of Computing Science and Engineering (SCSE), VIT Bhopal University, Sehore, India*
[5]*Artificial Intelligence and Data Science Department, Islamia Engineering College, Hyderabad, Telangana, India*
[6]*Department of CSE, University Centre for Research & Development, Chandigarh University, Mohali, Punjab 140413, India*
[7]*United International University, Dhaka, Bangladesh*

Correspondence should be addressed to Saima Ahmed Rahin; srahin213012@mscse.uiu.ac.bd

Academic Editor: Akshi Kumar

An environment of physically linked, technologically networked things that can be found online is known as the "Internet of Things." With the use of various devices connected to a network that allows data transfer between these devices, this includes the creation of intelligent communications and computational environments, such as intelligent homes, smart transportation systems, and intelligent FinTech. A variety of learning and optimization methods form the foundation of computational intelligence. Therefore, including new learning techniques such as opposition-based learning, optimization strategies, and reinforcement learning is the key growing trend for the next generation of IoT applications. In this study, a collaborative control system based on multiagent reinforcement learning with intelligent sensors for variable-guidance sections at various junctions is proposed. In the future generation of Internet of Things (IoT) applications, this study provides a multi-intersection variable steering lane-appropriate control approach that uses intelligent sensors to reduce traffic congestion at many junctions. Since the multi-intersection scene's complicated traffic flow cannot be accommodated by the conventional variable steering lane management approach. The priority experience replay algorithm is also included to improve the efficiency of the transition sequence's use in the experience replay pool and speed up the algorithm's convergence for effective quality of service in the upcoming IoT applications. The experimental investigation demonstrates that the multi-intersection variable steering lane with intelligent sensors is an appropriate control mechanism, successfully reducing queue length and delay time. The effectiveness of waiting times and other indicators is superior to that of other control methods, which efficiently coordinate the strategy switching of variable steerable lanes and enhance the traffic capacity of the road network under multiple intersections for effective quality of service in the upcoming IoT applications.

1. Introduction

With the continuous increase in the number of motor vehicles in my country, the contradiction between the supply and demand of road traffic is increasingly intensifying. Especially in the intersection scene, the traffic flow of each

turn at the intersection presents an uneven distribution at different time periods, which easily leads to congestion and waste of lanes resources. In order to solve this problem, the variable steerable lane technology using intelligent sensors came into being, which uses the lane as a variable space resource and dynamically allocates it according to the needs

of each turning traffic flow on the basis of communications between the wireless sensors, so as to improve the utilization rate of road space resources. An intelligent sensor is a function that can sense and detect information from a specific item, as well as learn, judge, and receive signals, and has a new type of sensor with management and communication features. The intelligent sensor is capable of autonomously calibrating, compensating, and collecting data. The intelligent sensor's capabilities determine its high precision and resolution, stability and dependability, and flexibility. It provides a great performance-to-price ratio when compared to standard sensors. Smart sensors come in three varieties: those that can judge; those that can learn; and those that can be creative. Intelligent velocity sensors, intelligent acceleration sensors, intelligent flow sensors, intelligent position sensors, intelligent attitude sensors, intelligent displacement sensors, and intelligent dimension sensors are a few examples of intelligent sensor types.

Variable steerable lanes at a single intersection of the scene, the traditional control method can effectively alleviate the problem of unbalanced steering, but as the number of variable steering lanes increases, the traffic flow changes between multiple intersections become more complex, and the present ability of the traditional method is more difficult, and it is impossible to coordinately control multiple. Therefore, how to make the cooperation between the multi-intersection variable-guidance lanes more efficient becomes a new problem. Variable steering lane with many intersections is a technique for figuring out the length of a changeable guide lane for a signal control junction approach is disclosed in an invention in the field of road traffic sign marked lines. The process entails first obtaining traffic data information on the variable guide lane's steering in the desired direction, followed by translating the received traffic volume into an equal standard small vehicle traffic volume. Using a model in the theory of queuing, it is then possible to determine the number of vehicles queuing behind a stop line for the signal control intersection approach using a multipath queuing multichannel system and a combination of an approach lane queuing imbalance coefficient. Finally, the vehicle number is converted to a measurement of the span of the variable guide lane for the signal control intersection approach using the vehicle average queuing length in meters. The approach traffic efficiency may be increased by using the technique for estimating the length of the variable guide lane for the signal control intersection approach, which allows cars on the upper stream of the variable guide lane to enter the intersection approach smoothly.

The term "variable-guided vehicle route" describes a track where the travel direction is not totally set for left-hand rotation, right turns, or continuing straight ahead, but instead incorporates various turning functions depending on the time of day. The overall traffic efficiency in the crossing can be greatly increased thanks to the way this track is laid out. Since the entrance driveway to the signalized crossing is not suitable for widening or having many canalizations, the wagon flow lacks uniformity significantly and is not suitable for the condition under which a signal control device can solve the problem. Apply in places like Hangzhou, Wuxi,

Tangshan, and Yantai, China's equivalent of Shanghai. Variable lanes are now being installed at crossroads in several places in order to better use road resources and increase traffic capacity. The variable lane control method primarily employs manual observation of traffic conditions for switching, or switching is performed at a fixed-time based on statistics of historical data. The manual switching mode is inefficient, and the statistical technique does not accurately reflect current traffic conditions. However, the rapid development of artificial intelligence and vehicle-road cooperation technologies can supply more real-time and precise data for variable lane control. It is feasible to regulate the variable lane more efficiently by making proper use of these data.

In this paper, we present a collaborative control system for variable-guidance sections at various intersections based on multiagent reinforcement learning and intelligent sensors. This study proposes a multi-intersection variable steering lane-appropriate control strategy that leverages intelligent sensors to minimize traffic congestion at multiple junctions in the next generation of Internet of Things (IoT) applications. The priority experience replay algorithm is also incorporated to increase the efficiency of the transition sequence's utilization in the experience replay pool and accelerate the algorithm's convergence for effective quality of service in forthcoming IoT applications.

2. Related Work

The traffic flow at the intersection changes dynamically in time and space. For example, in the morning and evening rush hours, the traffic flow of each turn at the intersection shows obvious regular changes, and there is a serious imbalance in the queue length of vehicles in different guide lanes. In order to improve the traffic capacity of the intersection and solve the problem of urban road congestion, some traffic researchers have carried out research on the dynamic control method of variable steering lane steering, mainly focusing on three aspects: traditional control method, intelligent control method, and reinforcement learning method. Supervised learning is a machine learning term that refers to the approach of constructing a function independently by learning from a number of related samples. This is the method of learning a broad notion from a small number of examples that are comparable to others. Contrarily, reinforcement learning is a branch of machine learning that builds on the idea of behavioral psychology and focuses on interacting with the environment directly. It is an important part of the field of artificial intelligence. Regression and classification are the two major jobs in supervised learning, whereas exploitation or exploration, Markov's decision processes, policy learning, deep learning, and value learning are the diverse tasks in reinforcement learning. Basic reinforcement is specified in the model Markov's decision process in reinforcement learning, whereas supervised learning examines the training data and generates a generalized formula. Each example in supervised learning will have a pair of input objects and an output with the desired values, whereas, in reinforcement learning, the

agent interacts with the environment in discrete steps, making an observation for each time period "t," receiving a reward for each observation, and then attempting to accumulate as many rewards as possible in order to make more observations.

The formula to measure the queuing length of the lanes can be explained as we can demonstrate for the M/M/1 queue that $L_q = \rho2/1 - \rho$

We can demonstrate for the M/G/1 queue that

$$L_q = \frac{\lambda^2 \sigma_s^2 + \rho^2}{2(1 - \rho)}, \tag{1}$$

where L_q is the mean number of customers in the queue, the likelihood that it is busy, and the percentage of time it is busy are all represented by the formula $\rho = \lambda/(c\mu)$, $1/E$ [Inter-arrival-Time], where E stands for the expectation operator, represents the mean rate of arrival, and σ_s^2 variance of service time.

2.1. Traditional Control Method.

The traditional variable lane control method research uses the method based on experience or historical data to set the control plan in advance and design the rules for the steering of the variable steerable lane at the intersection. Literature [1] proposed a signal for lane optimization at the intersection based on experience with phase-integrated design. Literature [2, 3] proposed a steering control model for a variable-steering lane at an intersection based on the empirical rules for the setting conditions of the variable-steering lane. The change characteristics of traffic flow and the actual traffic steering demand are more closely matched in Literature [4] and comprehensively considered the real-time traffic factors at a single intersection, and evaluated the preimplementation plan.

The specific lane function and signal timing switching scheme are deployed, but the preplan must be repeatedly tested and the accuracy is not high. Literature [5, 6] carried out integer nonlinear programming according to multiple road constraints associated with the target intersection. The optimization of the model achieves the smallest critical flow ratio after the optimization of a single intersection. Literature [7, 8] integrated the road conditions of key intersections and downstream adjacent intersections to achieve an associated control model. The above-mentioned work only considers the key intersection to a single adjacent intersection impact but does not design a comprehensive optimization scheme for associated intersections. Literature [9] proposed a control method to coordinate the design of multiple intersection variable signs and corresponding signal groups based on the collected data rules to better reduce the average delay of vehicles.

The abovementioned method of presetting the steering rules for variable steering lanes through experience or summarizing historical data rules can adapt to the needs of regular traffic state changes to a certain extent, but it is difficult to dynamically adapt to road traffic conditions and sudden changes in supply and demand abnormal traffic flow.

2.2. Intelligent Control Method.

Research on intelligent control methods for variable steering lanes makes intelligent decisions based on various traffic flow data collected in real-time, and improves the adaptability to real-time traffic flow changes at intersections. Some works use the collected real-time traffic flow data on roads, such as the space of each turning lane, occupancy rate [10], traffic flow, speed, queue length, and other characteristics obtained through video detection [11], dynamic decision-making variable steering lane switching strategy, but its adaptability to subsequent traffic flow changes is not good when combined with real-time collected data. Literature [12] predicted each turning traffic flow as the basis for judging lane direction switching and minimized the average delay time at the intersection. Literature [13] used the least squares dynamic weighting and the short-term traffic flow prediction model with fusion algorithm as the core combined with the traffic state prediction model with fuzzy data theory and neural network system theory as the core to realize automatic control of variable steering lane steering. Literature [14] constructed a mixed integer and the two-layer programming model is solved by a particle swarm algorithm to achieve the goal of minimizing the total travel time of variable lanes based on the prediction model.

The above-related research work has two limitations: (1) it is mainly applied to intelligent control decision-making of variable-steering lane steering at a single intersection and (2) the prediction-based intelligent algorithm is mainly based on historical and real-time data, and cannot quickly update rules to adapt to the dynamics of traffic flow variety.

2.3. Reinforcement Learning Methods.

In recent years, reinforcement learning technology has developed rapidly. It has low requirements for prior knowledge of the environment and can achieve good learning optimization performance in complex nonlinear systems. Therefore, it is suitable for complex and changeable multi-intersection and variable-guided lane intelligent control scenarios. In the multi-intersection collaborative control problem, traffic signal optimization research has widely used the reinforcement learning method. Literature [15] combined deep reinforcement learning with the traffic signal control problem, respectively, defining the state, action space, and reward function, using DQN (Deep Q-Network) model, extensive experiments on synthetic and real datasets demonstrate the superiority of reinforcement learning methods. Literature [16] used multiagent reinforcement learning technology to define the joint Q-value as the weighted sum of local Q-values, by minimizing the weighted sum of individual Q-values and the global Q-value to ensure that a single agent can take into account the learning process of other agents and realize automatic control of large-scale traffic signals. Literature [17] proposed that different agents exchange strategies after each round of learning to achieve a zero-sum game. Based on this realize the signal control strategy of autonomous vehicles, and designed a rewarding method that combines individual efficiency with overall efficiency. In the multi-intersection collaborative control scenario, the traffic

signal optimizes the traffic situation in the time dimension, and the intelligent variable guide lane is used as the spatial dimension. The above-mentioned two directions are suitable for the use of reinforcement learning methods to carry out global optimization research.

3. Variable Steering Lane Cooperative Control Method Employing Intelligent Sensors

3.1. Overall Structure. This research proposes a multi-intersection variable-direction lane cooperative control algorithm using intelligent sensors based on multiagent reinforcement learning. The method mainly includes multiagent reinforcement learning model merged with intelligent sensors, a global reward decomposition algorithm, and a priority experience playback algorithm. The multiagent reinforcement learning model proposed is constructed based on the value function decomposition algorithm of the QMIX algorithm [18]. The QMIX algorithm adopts the strategy of centralized training and distributed execution and uses the global reward function to optimize joint actions during training with the aid of intelligent sensors. The value function can achieve the effect of multiagent cooperative control, and each agent constructs and extracts the corresponding local strategy from the joint action-value function, which can not only deal with the problems caused by environmental nonstationary through centralized training but also through joint action-value function backpropagation learns the local "best" policy for each agent, enabling multiagent decentralized execution. Through the provision of a more flexible version of the constraint, QMIX enhances the VDN algorithm. The constraint is described as follows:

$$\frac{\partial \, \text{Qtot}}{\partial \, \text{Qa}} \geq , \tag{2}$$

where Q_{tot} denotes the total value function and Q_a denotes the value function for each agent. The weights of each particular value function Q_a should be positive, according to an obvious explanation. If the weights of the individual value function Q_a are negative, the agent will be discouraged from cooperating since the greater Q_a, the lower the joint value Q_{tot}. For consistency between the centralized and decentralized rules, QMIX comprises agent networks that represent each Q_a and a mixing network that brings them together into Q_{tot} rather than just adding them up like in VDN. By requiring the mixed network to have positive weights, it also imposes constraints. Since QMIX's factored representation grows well with the number of agents, it is able to describe complicated centralized action-value functions and makes it simple to extract decentralized policies using individual argmax operations in linear time.

The global reward decomposition algorithm improves the global reward distribution method in the value function decomposition algorithm and imposes constraints between the global value function and the value function of a single agent. In some complex scenarios, the global optimal joint action may require the intelligent sensor to make some behaviors that sacrifice individual interests. Decomposition

techniques change a difficult problem into a simpler one. Only binary constraints, whose scopes form a directed acyclic graph, are present in the new problem. Each set of variables from the original problem is represented by a variable in the new problem. These sets encompass the set of the initial variables even if they are not necessarily disjoint. In relation to each set of variables, the translation uncovers all partial solutions. The interplay between local solutions is reflected in the translation problem. A decomposition approach creates a binary acyclic issue by definition; these problems may be solved in a time that is polynomial in their size. In response to this problem, this study decomposes the global reward into two parts, one part is the basic reward, and the specific distribution to different agents is realized through the QMIX hybrid network; the other part is the performance reward, according to the agent. The state hierarchically is assigned to each agent which is the IoT so that a single agent can maximize the global reward while taking into account its own reward, and realize the secondary distribution of the global reward. In RL, the agent receives a reward that is often a sum of many reward components, each designed to encode some aspect of the desired agent behavior. From this composite reward, it learns a single composite value function. Using value decomposition, an agent learns a component value function for each reward component. To perform policy optimization, the composite value function is recovered by taking a weighted sum of the component value functions. While prior work has proposed value decomposition methods for discrete-action Q-learning. The development of autonomous agents is frequently done via reinforcement learning. In the RL framework, the agent is permitted to behave in the environment and is rewarded numerically at each step rather than being explicitly programmed. The goal of RL algorithms is to discover a strategy that maximizes the overall predicted reward (or some related criterion). Therefore, the reward function implies that optimal behavior is described. In order to assess actions in terms of trade-offs among the kinds, the technique decomposes rewards into sums of semantically relevant reward categories. To concisely describe why one action is preferable over another in terms of the kinds, we particularly propose the idea of minimal adequate explanations.

In the priority experience replay algorithm, in view of the uneven quality of randomly sampled experience, resulting in low training efficiency and slow algorithm convergence speed, the joint value function in the value function decomposition algorithm is used to calculate the error, and combined with the number of extractions to calculate priority of samples to speed up algorithm convergence.

3.2. Multiagent Sensors Reinforcement Learning Models. The multiagent sensor reinforcement learning model based on value function decomposition is shown in Figure 1. Based on the value-decomposition networks (VDN) algorithm [19], the original linear mapping is replaced by a nonlinear mapping, and the super network is introduced to add additional global state information to the mapping process to improve the performance of the algorithm [20]. Using the
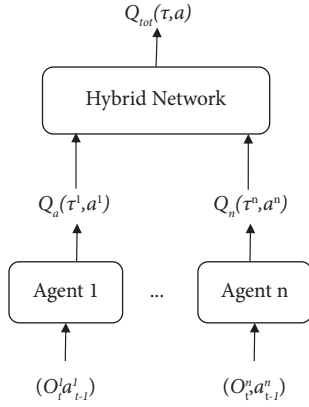
FIGURE 1: Multiagent reinforcement learning models.
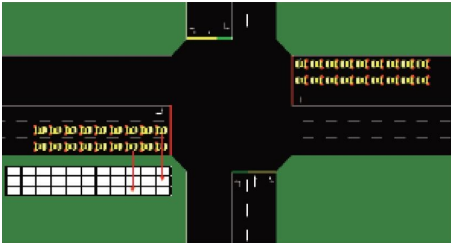
FIGURE 2: Schematic diagram of vehicle location matrix.

current observation state of each agent which are the intelligent sensors that performs the action of the previous time step as input, a global action-value function is learned through the hybrid network by these intelligent agents, where τ is the global state, α is the global action. In the multi-intersection variable steering lane scenario, the involved elements such as state space, action space, and reward function are defined. In order to make the input state more realistic and richer, the queue length of the lanes in each direction is adopted as the average waiting time of vehicles, and the ratio of average delay time as indicators. In addition, in order to accurately describe the position distribution of vehicles, the variable steering lane area is discretized and encoded to obtain the vehicle mapping matrix, as shown in Figure 2. The lanes are divided into the same size the grid covers the entire road section. Each grid in grid represents the present state of a vehicle. A value of 1 means that the vehicle exists at the grid, and 0 means that the vehicle does not exists compared with the intersection image directly information as input, this method can compress the data dimension and remove redundant information, thereby speeding up the training practice speed.

In the multi-intersection scenario, the state space expression is defined as follows:

$$S(T) = (L_L, W_L, D_L, L_s, W_S, D_S, M), \tag{3}$$

where T is the number of signal cycles, L_L, W_L, D_L are the average queue length, average waiting time, and average delay time of the lane group in the left-turn direction, respectively, L_s, W_S, D_S is the average queue length, average waiting time, and average delay time of the lane group in the

straight direction, and M is the vehicle position mapping matrix.

In the variable steering lane scenario, the straight-left variable steering lane is mainly studied and applied, and the right turn direction is not considered, so the action space is left turn or straight.

The global reward function is defined as the weighted sum of the following metrics.

(1) The average queue length L of vehicles in all lanes

(2) The average delay time ratio D of vehicles in all lanes

The expression of the single lane delay time ratio D_i is as follows:

$$D_i = 1 - \frac{v_{lane}}{v_{max}}, \tag{4}$$

where v_{lane} is the average speed of vehicles on lane i, and v_{max} is the vehicle speed maximum speed limit for road i:

(3) The average waiting time W of vehicles in all lanes. When a vehicle starts to stop and wait, that is, when the speed is less than 0.1 m/s, the parking waiting time of this vehicle starts to accumulate;

(4) N is the average number of vehicles on all lanes that left the current lane after the previous action;

(5) V is the average speed of vehicles in all lanes leaving the current lane after the previous action. The expression of the average speed is given as follows:

$$V = \frac{1}{N} \sum_{j=1}^{N} v_j, \tag{5}$$

where v_j is the average speed of each vehicle.

Assign corresponding weights to the above-mentioned different traffic indicators, and finally calculate the global reward:

$$R = k_1 L + k_2 D + k_3 W + k_4 N + k_5 V. \tag{6}$$

In the formula, $k1, k2, k3, k4, k5$ are the weight parameters, and the effect of the final traffic condition optimization is set by analyzing the global reward results in 3200 experiments.

3.3. Global Reward Decomposition Algorithm.

The global reward R is decomposed into two parts, the basic reward R_b and the performance reward R_p based on the proportion, as shown in Figure 3 for the global reward decomposition algorithm structure. The performance reward is an additional reward that is used to distribute to the agents with larger contributions.

The global reward decomposition function is given as follows:

$$R = R_b + R_p,$$
$$R_b = (1 - \lambda)R, R_p = \lambda R; \lambda \in [0, 1.0]. \tag{7}$$
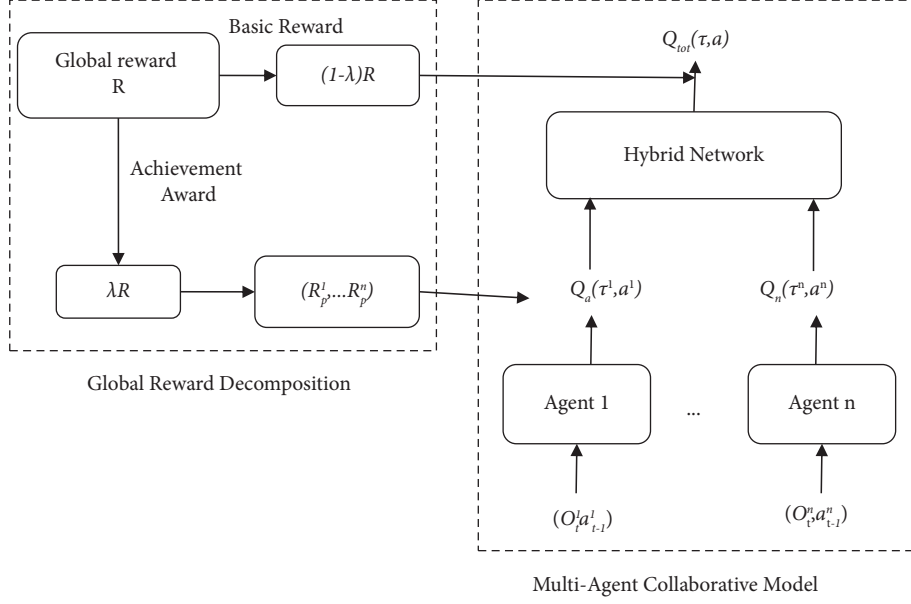
FIGURE 3: The global reward decomposition algorithm.

The traditional hybrid network method is used to distribute the basic rewards to each agent which are the intelligent sensors. The performance reward is used to motivate the agents that contribute more in the process of regional cooperative control. The expression of the performance reward obtained by each agent at the current time is given as follows:

Real-world situations like strategic conflict resolution, coordination between autonomous vehicles, and agent cooperation in defensive escort squads all include cooperative multiagent challenges. Such issues may be modeled as dual-interest situations, in which each agent simultaneously seeks to maximize its individual payout (local reward) and the team's performance as a whole (global reward). Two different forms of modern, state-of-the-art MARL algorithms exist. While algorithms like MADDPG and M3DDPG concentrate on optimizing local rewards without any explicit idea of coordination, algorithms like COMA and QMIX strive to maximize the global reward for the success of the group. We first define multiagent cooperation as a joint optimization on reward assignment and demonstrate that each agent has an approximately optimum strategy that decomposes into two parts: one that just depends on the agent's own state and the other that is connected to the states of adjacent agents. CollaQ decomposes each agent's Q-function into a self-term and an interacting term, using a multiagent reward attribution (MARA) loss to regularize training. CollaQ is tested on multiple StarCraft maps and surpasses existing state-of-the-art approaches (such as QMIX, QTRAN, and VDN) by increasing the win rate by 40% while using the same amount of samples.

$$
R_p^i = \begin{cases} \dfrac{1}{2L_T}\left(L_{S'} + L_{L'}\right)R_p, & L_{S'} < L_T, L_{L'} < L_T, \\[2mm] R_p, & other, \end{cases}
$$

$$
L_T = \frac{V_{out}}{V_{\max}},
$$

$$
L' = \frac{L}{\overline{L}}.
$$

(8)

In the formula, R_p^i is the performance reward obtained by the ith agent, L' is the ratio of the average queuing length of the lane group in a certain direction of the agent to the overall length of the lane, $\overline{L}$ is the average queuing length of the lane group during the execution of the previous decision, $L_{S'}$ is the overall length of the lane, and $L_{L'}$ is the ratio of the average queuing length of the straight lane group to the overall length of the lane, the ratio of the average queuing length of the left-turn lane group, L_T the threshold of the determination level of the lane, V_{out} the maximum traffic flow that can be driven out during the green light of the lane, and the capacity of the lane. $V_{\max}$ is the maximum traffic flow [21].

All agents are graded when the performance reward is allocated R_p, and the performance reward corresponding to each level is different [22, 23]. As shown in formula (8), when the $L_{L'}$ left turn and $L_{S'}$ straight lane group queue length ratio and are L_T less than the threshold when the traffic flow of the road section is determined to be small, $L_{L'}$ and $L_{s'}$ the performance reward is allocated by the ratio of

the average and the larger, more performance rewards should be allocated [24].

3.4. Priority Experience Replay Algorithm.

Deep learning uses target networks to increase the training's stability. The main training network and the target network are the two networks that the DQN method trains. The difference between the two networks, squared, is the loss the algorithm trains on (often replaced with Huber loss nowadays). The primary training network periodically replaces the target network's weights as training progresses. The target network predicts the optimum Q value from all actions that may be done from the next state in each data sample. This is the desired Q value. To train the Q network, the loss is computed using the predicted Q value, target Q value, and observed reward from the data sample. In single-agent reinforcement learning, in order to solve the problem of uneven quality of training samples extracted during the training process, a priority experience playback algorithm [25] was proposed, and the temporal-difference (TD) method was used to measure the importance of samples. The samples with larger errors are set as high priority, and the samples with high priority are extracted for training to improve learning efficiency. In the multiagent where agents are the intelligent sensors the reinforcement learning based on the value function decomposition algorithm, the joint value function can be used to calculate the TD error, and then use it for the calculation of the priority. In order to realize the priority experience playback algorithm, the target network loss L_n must be calculated:

$$
\begin{aligned}
L_n &= \left(y^{tot} - Q_{tot}\,(S, A) \right)^2, \\
y^{tot} &= R + \gamma Q_{tot}\,(S', A').
\end{aligned}
\tag{9}
$$

In the formula, S, A are the joint state and joint action of multiagent, γ is the attenuation coefficient, S', A' is the joint state and joint action of multiagent at the next moment. The larger the value, L_n the higher the corresponding experience sequence will be priority.

Using L_n as the only indicator to measure the importance of samples may cause some samples to be drawn less frequently due to their small size. Therefore, this study combines the target network loss and the number of times to be drawn N_{sam} as an indicator to measure the importance of samples. At the same time, considering different empirical the loss value of the target network has a large difference $f_{Top}(L_n)$, and it is converted into a dimensionless sorting amount, which is the position of the loss in the increasing sorting. The expression of the final priority PR_i is given as follows:

$$
\begin{aligned}
PR_i &= \frac{p_i}{\sum_{j=1}^M p_j + \epsilon p_i} \\
&= f_{Top}\left(f_{Top}\left(L_{ni}\right) + f_{Bot}\left(N_{sam}^i\right) \right).
\end{aligned}
\tag{10}
$$

In the formula, $f_{Bot}(N_{sam}^i)$ is the position of the number of extractions in descending order; $\varepsilon \in (0, 1.0)$ is the offset of the probability, which is used to correct the situation that the priority is too small to cause the probability of sample selection to be too low.

4. Experimental Results and Analysis

In order to verify the effectiveness of the cooperative control algorithm that employs intelligent sensors to alleviate the traffic congestion at multiple intersections methods for efficient quality of service in the next generation IoT applications using reinforcement learning. In the multi-intersection variable steerable lane scenario, the cooperative control BASE algorithm is combined with fixed-time control (FT) and traditional adaptive control algorithm for multi-intersection (MTAC), single-agent reinforcement learning adaptive algorithm (DQN), multiagent reinforcement learning adaptive algorithm (QMIX) and other methods, and analyze the performance of each algorithm in the data set, including algorithm-level reward value, traffic average queue length, average delay-to-time ratio, average waiting time, and average travel time metrics at the level. In cooperative multiagent systems, agents work together to complete tasks in exchange for a group reward as opposed to individual benefits. Credit assignment techniques are often used to differentiate the contributions of various agents in the absence of individual reward signals in order to promote successful cooperation. As credit assignment has recently been widely implemented using the value decomposition paradigm, QMIX has emerged as the leading technology. Robot swarms, autonomous vehicles, sensor networks, and cooperative multiagent reinforcement learning are just a few areas where this technique has found widespread use. In these activities, each agent must learn a decentralized policy through a shared team reward signal because individual incentives are not available. In order to achieve successful collaboration, agents must allocate credit in a discriminatory manner. Credit assignment techniques using cooperative multiagent reinforcement learning have made significant strides in recent years. Value-based approaches have among them demonstrated cutting-edge performance on difficult challenges. Value factorization, which is based on the centralized training with decentralized execution (CTDE) paradigm, has recently gained a lot of popularity. It specifically integrates separate value functions Q_i during centralized training to factorize the combined value function Q_{tot}. Decentralized policies may be easily determined during execution by greedily picking individual actions from the local value function Q_i. Because Q_i is learned by maximizing the overall temporal-difference error on a single global reward signal, an implicit multiagent credit assignment is accomplished.

4.1. Experimental Setup.

The experimental equipment configuration of this research is as follows: the CPU is AMD 2.10 GHz, the memory is 16 GB, and the operating system is Windows 10 (64 bit). Simulation experiments are carried out based on the microscopic traffic simulation platform SUMO v1.7.0. Interface to interact with the simulation

Real Picture of Road Network   Experimental Area Road
in Experimental Area   Network Simulation Structure
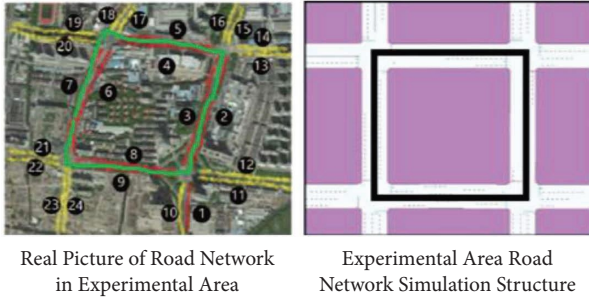
FIGURE 4: Structure diagram of the experimental road network area.

environment, obtain the traffic status in real-time and adaptively adjust the control strategy of the variable steering lane. As shown in Figure 4, the experimental environment includes 4 intersections and a total of 24 road sections, and the road sections are numbered 1~9. There are 9 variable steering lanes in total. The section with variable steering lanes consists of 5 lanes: fixed left turn, variable left turn/ straight, straight, straight, and right turn lanes. Section number 10~24 is a conventional road section, there are 15 in total, and the road section adopts a conventional fixed three-lane configuration. The signal cycle of each intersection is the same.

The experimental data set is collected from different traffic capture data of streets, roads, etc. such as upstream and downstream road section codes, capture time, lane number, and license plate number.

In this experiment, traffic flow data were collected at intersections in the urban area. The types of vehicles mainly included cars, station wagons, buses, and large passenger cars. The actual collected vehicle types and numbers were input into the simulation system, which were 24,17,592 cars, 156,008 station wagons, 91,089 buses, and 25,113 large passenger cars, accounting for 89.88%, 5.80%, 3.39%, and 0.93% of the total flow, respectively. In the calculation process of the simulation system, in order to standardize the vehicle position information at the intersection is converted into a discretized encoded vehicle position matrix, which is used as the quantitative input of the reinforcement learning model. In addition, the actual input vehicle type and quantity are converted into standard cars based on the conversion standard. The equivalent conversion coefficients of the models corresponding to standard cars are 1.2 for station wagons, 2.0 for buses, and large passenger cars. As is well known, traffic noise conversion and calculation used to be done using the prediction model in "Specifications for the Environmental Impact Assessment of Highways," whereas we now know that its levels or grades, as well as its equivalent conversion, can be calculated in a variety of ways using the equivalent conversion coefficient. The findings of the speed survey and their analysis allow us to conclude that the speed computation for all types of vehicles must follow the Gaussian distribution in free traffic flow, allowing us to suggest a speed discretization approach. The approach described above assists in converting various vehicles with varying speeds into car comparable numbers, while overall

traffic volume may be converted into that of passenger cars by evaluating themselves at the same noise levels.

In order to ensure the fairness of the comparison algorithm, the network structure and hyperparameter settings of the reinforcement learning algorithm are the same. The value of the discount factor is 0.95, the value of the learning rate is 0.001, the value of the greedy strategy ε is 0.05, and the size of the memory bank is 1000. The number of samples for each update is 32, and the model update step is 5 signal cycles. In order to improve the stability of the algorithm, the target network is updated with a delay. The weight replacement step of the target network is set to 30 signal cycles and RMS prop (root mean square prop) algorithm is the update algorithm.

In this study, the weight of global reward decomposition is set to $\lambda = 0.4$ in the BASE algorithm. Queuing length is used as an indicator to measure the efficiency of road traffic diversion under unbalanced traffic flow, and 32 combinations of weight values are set. In the process of training and testing, the accumulated vehicle queue length is calculated, 100 experiments are carried out for each scheme, and the average queue length is obtained by taking the average of the results. The smaller the average queue length the better the traffic diversion effect a total of 3,200 experiments were conducted using the numerous sensors. The results determine the weight of each influencing factor in the global reward function set in this experiment $k1, k2, k3, k4, k5$ are set to $-1.0, -0.5, -0.5, 0.5, 1.0$, respectively.

4.2. Experiment Analysis. Performance the test set is data of 6 periods, namely morning peak (periods 1, 2), evening peak (periods 3, 4), and flat peak (periods 5, 6). The performance of each control method is shown in Figures 5 and 6. In Figure 5, R is the reward index. In the data sets of multiple periods, the BASE algorithm performs better overall, followed by the QMIX algorithm, indicating that the multi-agent collaborative algorithm is still effective in real scenarios. The performance of BASE, QMIX, IQL, and MTAC is significantly better than that of the timing control scheme (FT), indicating that the adaptive algorithm can effectively adapt to the traffic flow changes in the real scene, and can make appropriate decisions according to the real-time traffic state.

Figure 5 and Table 1 show that the performance of the BASE algorithm in the datasets of multiple time periods is better than that of other algorithms, and the BASE algorithm has a significant lead in the morning and evening peak hours compared to the flat peak hours.

As shown in Figure 6 and Table 2, compared with other algorithms, the average queuing length index of the BASE algorithm in the morning and evening peak hours is reduced by 25.76% ~ 54.97%, and the index in the off-peak hours is reduced by 49.00% ~ 70.67%, indicating that the BASE algorithm is in the congestion scene. Compared with other algorithms, the average delay time of the BASE algorithm is reduced by 15.54%~55.09% as shown in Figure 7 and Table 3. In the test sets of 5 and 6 during the peak period, the road traffic in the network is small, the demand for the
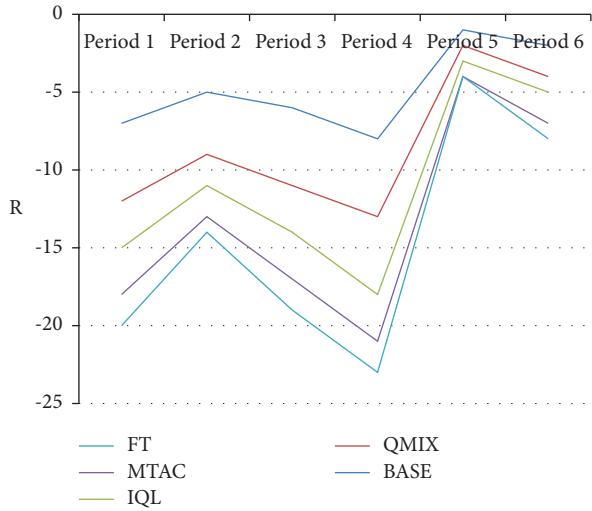
Figure 5: Comparison results of algorithm reward indicators in the test set.

Table 1: Comparison results of algorithm reward indicators in the test set.

BASE	QMIX	IQL	MTAC	FT
−7	−5	−3	−3	−2
−5	−4	−2	−2	−1
−6	−5	−3	−3	−2
−8	−5	−5	−3	−2
−1	−1	−1	−1	0
−2	−2	−1	−2	−1

Table 2: Average queue length.

BASE	QMIX	IQL	MTAC	FT
18	24	30	34	42
17	25	32	31	44
19	24	27	32	41
20	27	31	33	45
2	3	3	4	4
4	5	5	5	6
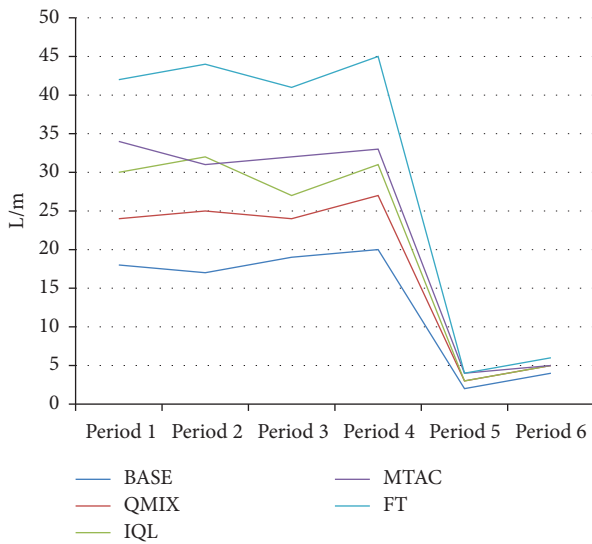
Figure 6: Average queue length.
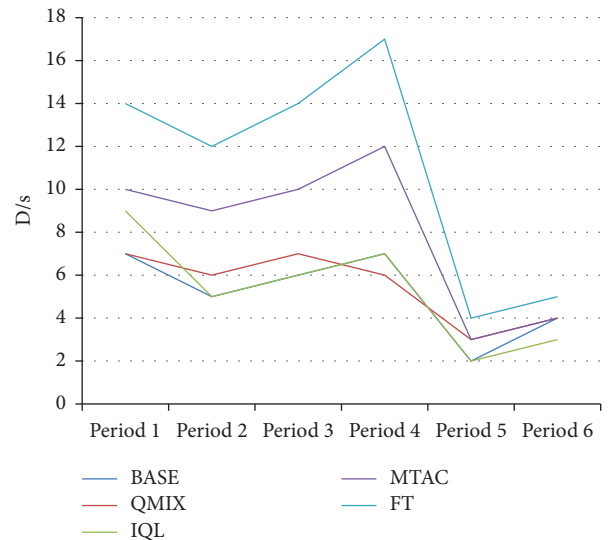
Figure 7: Average delay time.

function of the variable steering lane is weak, and the performance of each algorithm is close. The performance of the algorithm (BASE) in this study still maintains a slight lead. As shown in Figure 8 and Table 4, compared with other algorithms, in the BASE algorithm the average waiting time is reduced by 9.28%~42.39%, and in peak hours such as time period 3, the traffic state is more congested. The average waiting time of the IQL algorithm on this dataset is slightly better than that of the QMIX algorithm.

It can still maintain the leading performance, which further proves the effectiveness of the improved algorithm. As shown in Figure 9 and Table 5, the average travel time is reduced by 6.44%~29.93% compared with other algorithms, and the performance is stable under the test set of 6 periods, always has better performance.

Comparison results of each traffic index in the test set are shown in Figures 6 to 9.

The best performance of the algorithm in this study verifies the improvement of the multiagent collaborative algorithm for the multi-intersection variable-guidance lane scene: global reward decomposition, the proposed algorithm in this study can learn the policy better than the QMIX algorithm performance of the training process by comparing the BASE algorithm with the QMIX algorithm, the performance indicators of the training process are tested. The comparison of the average accumulated reward value index is shown in Figure – and Table 6. In Figure 10, E is the number of iterations.

The changes in traffic indicators during the training process are shown in Figure 8. Traffic indicators such as

TABLE 3: Average delay time.

BASE	QMIX	IQL	MTAC	FT
7	7	9	10	14
5	6	5	9	12
6	7	6	10	14
7	6	7	12	17
2	3	2	3	4
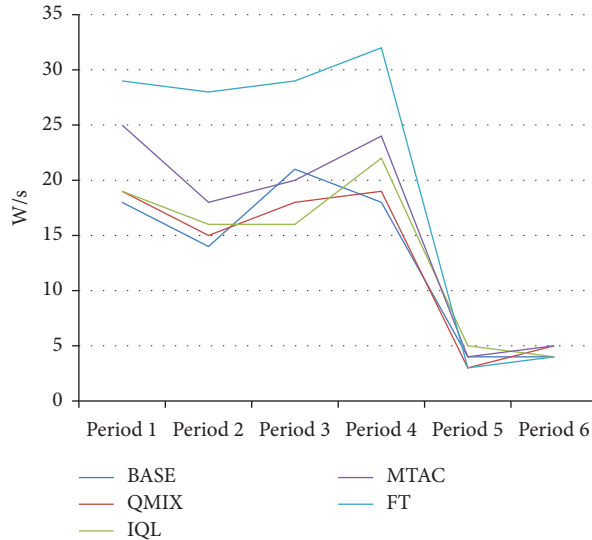
FIGURE 8: Average waiting time.

TABLE 4: Average waiting time.

BASE	QMIX	IQL	MTAC	FT
18	19	19	25	29
14	15	16	18	28
21	18	16	20	29
18	19	22	24	32
4	3	5	4	3

TABLE 5: Average travel time.

BASE	QMIX	IQL	MTAC	FT
41	43	44	48	54
38	41	43	45	51
40	42	44	44	62
39	41	43	48	60
19	20	24	25	28

TABLE 6: Average accumulated reward value index.

Average queue length		Average delay time		Average waiting time		Average travel time	
BASE	QMIX	BASE	QMIX	BASE	QMIX	BASE	QMIX
59	58	39	38	60	59	64	68
21	37	6	15	19	35	48	58
17	28	5	12	18	28	44	55
18	25	4	8	18	23	43	51
18	23	4	7	18	20	44	47
17	21	3	9	17	21	45	49
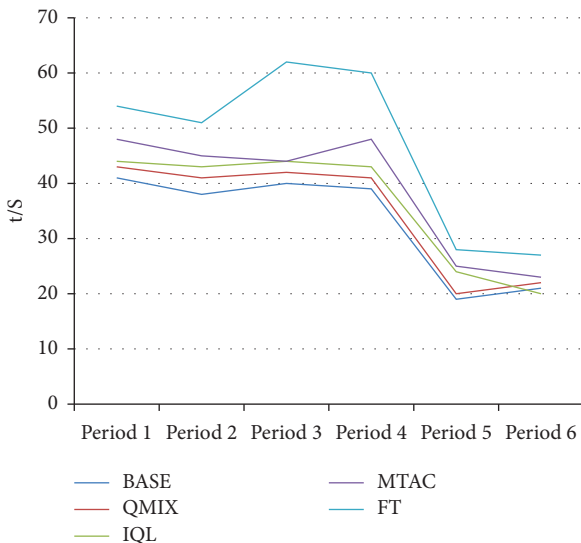
FIGURE 9: Average travel time.

average queue length, average delay time, average waiting time, and average travel time change with the training process of the algorithm model. The indicators are significantly reduced, the traffic state is gradually optimized, and the two multiagent collaborative algorithms can converge. At the same time, according to the downward trend of each traffic indicator in the training process and the final convergence state, it can be found that the global reward decomposition and priority experience playback are applied. The BASE algorithm of
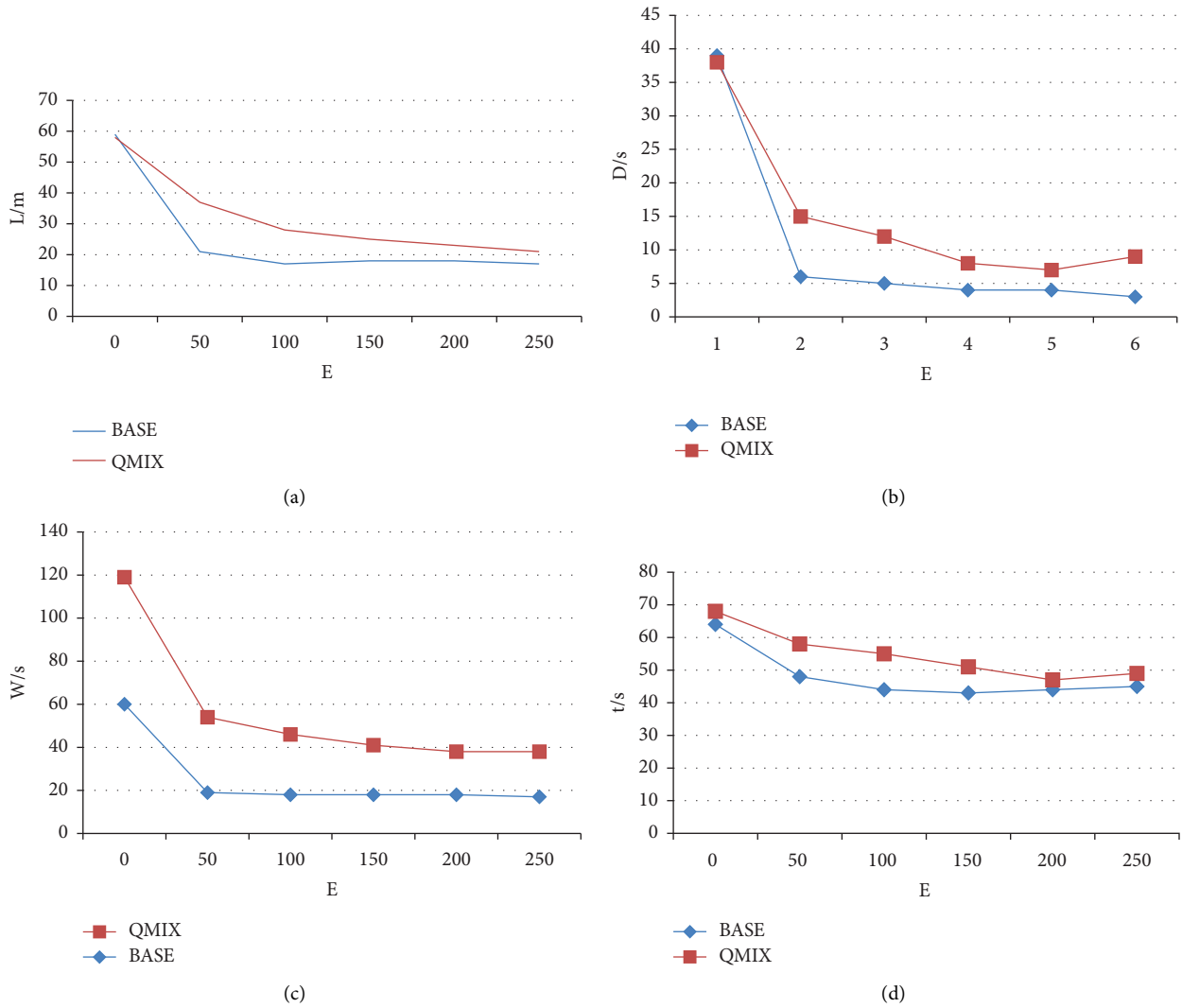
Figure 10: (a): Average queue length. (b): Average delay time. (c): Average waiting time. (d): Average travel time.

algorithm has a faster convergence effect and better optimization performance, which can effectively adapt to the traffic flow of multi-intersection scenarios, and realize traffic state optimization of the local road network.

5. Conclusion

Computational intelligence is built on several learning and optimization techniques. Therefore, an emerging trend of major significance in the future generation of IoT applications is the integration of new learning methods, such as opposition-based learning, optimization approaches, and reinforcement learning. This research proposes a multiagent reinforcement learning-based collaborative control method employing intelligent sensors for variable-guidance lanes at multiple intersections. This method improves the performance in congested scenarios through a global reward decomposition algorithm and improves learning efficiency through a priority experience replay algorithm. Cooperative control of variable-guidance lanes at multiple intersections: compared with other control methods, this method has better effects in reducing the average queue length, average delay time, average travel time, etc., while convergent faster, and therefore, enhancing the quality of service in the next generation of IoT applications.

The follow-up work includes combining the algorithm with traffic signal control and performing joint optimization in the two dimensions of time and space to further improve the traffic capacity of multi-intersection scenarios.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. Do and C. Lee, "Multi-agent reinforcement learning in a large scale environment via supervisory network and curriculum learning," in *Proceedings of the 2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 207–210, Jeju, South Korea, October 2021.

[2] M. Ye, C. Tianqing, and F. Wenhui, "A single-task and multi-decision evolutionary game model based on multi-agent reinforcement learning," *Journal of Systems Engineering and Electronics*, vol. 32, no. 3, pp. 642–657, 2021.

[3] J. Rosenberger, M. Urlaub, and D. Schramm, "Multi-agent reinforcement learning for intelligent resource allocation in IIoT networks," in *Proceedings of the 2021 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)*, pp. 118-119, Dubai, UAE, December 2021.

[4] J. Huang, Z. Ji, S. Xiao, C. Jia, Y. Jia, and X. Wang, "Multi-Agent Vehicle Formation Control Based on MPC and Particle Swarm Optimization Algorithm," in *Proceedings of the 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 288–292, Chongqing, China, March 2022.

[5] R. T. van Katwijk, B. De Schutter, and J. Hellendoorn, "Multi-agent control of traffic networks: algorithm and case study," in *Proceedings of the 2009 12th International IEEE Conference on Intelligent Transportation Systems*, pp. 1–6, St. Louis, Missouri, October 2009.

[6] B. Park, C. Kang, and J. Choi, "Generating multi-agent patrol areas by reinforcement learning," in *Proceedings of the 2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 104–107, Jeju, Korea, October 2021.

[7] S. Yeo, S. Lee, B. Choi, and S. Oh, "Integrate multi-agent simulation environment and multi-agent reinforcement learning (MARL) for real-world scenario," in *Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 523–525, Jeju, Korea, October 2020.

[8] F. Gao, S. Chen, M. Li, and B. Huang, "MaCA: A Multi-Agent Reinforcement Learning Platform for Collective Intelligence," in *Proceedings of the 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS)*, pp. 108–111, China, October 2019.

[9] E. D. Ferreira, E. Subrahmanian, and D. Manstetten, "Intelligent agents in decentralized traffic control," in *Proceedings of the ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)*, pp. 705–709, Oakland, CA, USA, August 2001.

[10] S. Jiang, Y. Huang, M. Jafari, and M. Jalayer, "A distributed multi-agent reinforcement learning with graph decomposition approach for large-scale Adaptive traffic signal control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14689–14701, 2022.

[11] H. Ge, D. Gao, L. Sun et al., "Multi-agent transfer reinforcement learning with multi-view encoder for adaptive traffic signal control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12572–12587, 2022.

[12] M. P. Lokhande, D. D. Patil, L. V. Patil, and M. Shabaz, "Machine-to-Machine communication for device identification and classification in secure telerobotics surgery," in *Security and Communication Networks*, C. Chakraborty, Ed., vol. 2021, Article ID 5287514, pp. 1-16, 2021.

[13] V. Mahalingam and A. Agrawal, "Learning agents based intelligent transport and routing systems for autonomous vehicles and their respective vehicle control systems based on model predictive control (MPC)," in *Proceedings of the 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pp. 284–290, Bangalore, India, May 2016.

[14] C.-G. Li, M. Wang, and Q.-N. Yuan, "A Multi-Agent Reinforcement Learning Using Actor-Critic Methods," in *Proceedings of the 2008 International Conference on Machine Learning and Cybernetics*, pp. 878–882, Kunming, China, July 2008.

[15] Z. Jiandong, Y. Qiming, S. Guoqing, L. Yi, and W. Yong, "UAV cooperative air combat maneuver decision based on multi-agent reinforcement learning," *Journal of Systems Engineering and Electronics*, vol. 32, no. 6, pp. 1421–1438, Dec, 2021.

[16] A. Kumar and D. Gupta, "Paradigm shift from conventional software quality models to web based quality models," *International Journal of Hybrid Intelligent Systems*, vol. 14, no. 3, pp. 167–179, 2018.

[17] N. Verma, M. Rakhra, M. W. Bhatt, and U. Garg, "Engineering technology characterization of source solution for ZnO and their data analytics effect with aloe vera extract," *Neuroscience Informatics*, vol. 2, no. Issue 3, Article ID 100015, 2022.

[18] Y. Lei, S. Vyas, S. Gupta, and M. Shabaz, "AI based study on product development and process design," *International Journal of System Assurance Engineering and Management*, vol. 13, pp. 305–311, 2021.

[19] G. Hu, Y. Zhu, D. Zhao, M. Zhao, and J. Hao, "Event-triggered communication network with limited-bandwidth constraint for multi-agent reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–13, 2021.

[20] A. Kumar and A. Arora, "Website quality analytics using metaheuristic based optimization," *Recent Advances in Computer Science and Communications*, vol. 14, no. 3, pp. 895–915, 2021.

[21] X. Fang, Q. Zhao, J. Wang, Y. Han, and Y. Li, "Multi-agent deep reinforcement learning for distributed energy management and strategy optimization of microgrid market," *Sustainable Cities and Society*, vol. 74, Article ID 103163, 2021.

[22] H. Zhang, J. Liu, Z. Wang, H. Yan, and C. Zhang, "Distributed adaptive event-triggered control and stability analysis for vehicular platoon," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1627–1638, 2021.

[23] L. Han, S. Su, D. Li, and D. Wang, "Finite-Time Distributed Adaptive Coordinated Control for Multiple Electric Units of High Speed Train," in *Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 1119–1124, Indianapolis, IN, USA, September 2021.

[24] S. El-Tantawy and B. Abdulhai, "Multi-Agent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC)," in *Proceedings of the 2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 319–326, Alaska, USA, September 2012.

[25] J. Zhang, L. Yan, Y. Han, G. Song, and X. Fang, "Research on the Method and Simulation of Intersection Signal Control Based on Multi-Agent," in *Proceedings of the 2009 International Conference on Management and Service Science*, pp. 1–4, Beijing, China, September 2009.

Research Article

Acceleration of Deep Neural Network Training Using Field Programmable Gate Arrays

Guta Tesema Tufa,[1] **Fitsum Assamnew Andargie,**[2] **and Anchit Bijalwan** [ID][3]

[1]*Faculty of Electrical and Computer Engineering, Arba Minch Institute of Technology, Arba Minch, Ethiopia*
[2]*School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Ethiopia*
[3]*School of Computing and Innovative Technologies, British University Vietnam, Hu'ng Yên, Vietnam*

Correspondence should be addressed to Anchit Bijalwan; anchit.bijalwan@amu.edu.et

Convolutional neural network (CNN) training often necessitates a considerable amount of computational resources. In recent years, several studies have proposed for CNN inference and training accelerators in which the FPGAs have previously demonstrated good performance and energy efficiency. To speed up the processing, CNN requires additional computational resources such as memory bandwidth, a FPGA platform resource usage, time, power consumption, and large datasets for training. They are constrained by the requirement for improved hardware acceleration to support scalability beyond existing data and model sizes. This paper proposes a procedure for energy efficient CNN training in collaboration with an FPGA-based accelerator. We employed optimizations such as quantization, which is a common model compression technique, to speed up the CNN training process. Additionally, a gradient accumulation buffer is used to ensure maximum operating efficiency while maintaining gradient descent of the learning algorithm. To validate the design, we implemented the AlexNet and VGG-16 models on an FPGA board and laptop CPU along side GPU. It achieves 203.75 GOPS on Terasic DE1 SoC with the AlexNet model and 196.50 GOPS with the VGG-16 model on Terasic DE-SoC. Our result also exhibits that the FPGA accelerators are more energy efficient than other platforms.

1. Introduction

In recent years, deep learning has shown their usefulness and effectiveness in finding an answer to many actual world problems. The DNN, notably the convolutional neural network, is at the root of this renaissance. The convolution neural network was shown to be a useful tool for a variety of functions, including image classification [1], image recognition [2], and object detection [3]. A CNN involves a massive number of computations that it could profit from acceleration using GPUs and FPGAs [4, 5]. Deep CNN hardware implementations are constrained by a memory bottleneck that need numerous convolutions and fully connected layers, which necessitate a considerable amount of communication for parallel processing [6].

A variety of accelerators, including graphics processing units (GPUs), Field Programmable Gate Arrays, and application specific integrated circuits, has been used to increase the efficiency of CNNs [7–9]. Among these accelerators, GPUs are the most commonly employed to enhance throughput and memory bandwidth [8], both in the training and the inference process of CNN; however, they use high power [1, 6, 10]. An alternatively, field programmable gate arrays (FPGAs) are a natural option for neural network deployment since computing, logic, and memory resources may be merged into a single device. Based on FPGAs (field programmable gate arrays), CNN accelerators provide significant benefits because of their reduced power consumption, high throughput, and design flexibility [11]. FPGAs also provide high parallelism and exploit the features of neural network processing [12]. However, CNN on FPGA

has a number of challenges such as requirements of memory storage, external memory bandwidth, and computational resource limitations. However, the FPGA restricted resources, such as the Stratix A7, have close effects to the midrange FPGA (Arria GX 10) citeli2017acceleration. The previous hardware accelerators for CNN have used different kernel for convolution and fully connected layers, which affect the FPGAs resource utilization [5, 13].

Intel's programmable solutions division has created a scalable convolutional neural network reference architecture for deep learning systems based on the OpenCL programming language. The OpenCL-based design tool is used to effectively accomplish the required accelerator design. This allows us to reuse the current code for Graphics Processing Units (GPUs) in FPGAs using OpenCL-based high-level synthesis tools [6, 14]. Developers may program the FPGAs in high-level languages like as C/C++ using high-level synthesis (HLS), which speeds up the development process. HLS techniques provide a developer with an extremely simple programming model as FPGA [12]. However, the CNNs are mostly solved using methods based on matrix-multiplication; this somehow requires the movement of huge volumes of data between compute units and external memory [5]. To speed up processing, the CNN requires more computational resources. Nonetheless, when processing CNNs, a memory bandwidth is often the bottleneck. Because of the high memory requirements of the fully connected (FC) layers, layer sections and the execution might be memory limited. The enormous number of weights held by these layers accounts for the high number of memory reads. If any of these accesses are to external memory, for instance, dynamic random access memory, throughput and energy power usage would be significantly impacted because dynamic random access memory accesses have far more latency and energy consumption than the compute itself.

However, memory storage, external memory bandwidth, and computing resource limits provide a number of challenges for CNN on FPGA.

The contributions of this work are as follows:

(1) It proposes single kernel for both convolutional and FC layers, which improve memory bandwidth and hardware resource utilization.

(2) Loop parallelization and single instruction multiple data (SIMD) have been applied.

(3) To get maximum throughput, we use design space exploration method that leverages resource usage and throughput and is able to find the optimal architecture configuration, for CNN on FPGA.

2. Background

This section explains the basic theoretical basis for solving image classification problems. As such, we explain how hardware accelerators are used for image classification by first giving brief description of the hardware platforms and convolution neural network.

2.1. FPGA Architecture. FPGAs (Field Programmable Gate Arrays) were first used nearly two and a half decades ago. FPGAs are semiconductor devices that are built around a grid of configurable logic blocks (CLBs) interlinked via programmable interconnects. The FPGAs are programmable devices that offer a versatile platform for developing unique hardware capabilities at a reduced development cost [15]. The modern FPGA has two main parts: programmable logic blocks (ALMs) and logic components [12]. Figure 1 shows that the FPGA has a different configurable logic block (CLB) as well as input and output ports. The configurable logic block (CLB) is the basic repeating logic resource on an FPGA, which contains smaller components, such as flip-flops, look-up tables (LUTs), and multiplexers. The FPGA resources that allow connecting the FPGA target to other devices are the input and output (I/O). Input and output are to change analog or digital signals to or from a digital value so that we can process the signals using an FPGA target. The FPGAs logic capacity has been greatly increased because of advancements in process technology, making them a feasible implementation option for bigger and more sophisticated designs. Generally, the FPGA nature of logic and resource usage affects the FPGA device's space, speed, and power efficiency [16].

2.2. Intel FPGA SDK for OpenCL. A high-level abstraction for FPGA programming is provided by the Intel OpenCL SDK as one of the HLS tools. A concurrent program is built to von Neumann fixed structure as shown in a series of instructions for hardware acceleration that each computation generally requires the retrieval of instructions as well as the moving of data between register data and also the memory [17]. The Intel OpenCL SDK solution, on the other hand, provides a highly effective solution. Inside this model, the platform resources are customized to the algorithm being run [17].

Global memory is arranged as external memory in the FPGA device for the memory system in the Intel OpenCL SDK, which could be DDR3 synchronous dynamic random access memory as well as other memory [18].

2.3. Convolutional Neural Networks. CNN is a type of deep neural network that is very useful for classification. It takes an input and predicts a class tag for it. CNN typically consists of many layers, such as convolutional layers, ReLU layers, pooling layers, normalization layers, and fully connected layers. So, every layer will have its own input and output, with the input mapped to either a linear or nonlinear transformation of the output. Below are listed a descriptions of the individual layers.

2.3.1. Convolutional Layer. The convolution layer parameters are made up of a series of learnable filters. Each filter has a small spatial footprint, but extending to the maximum depth of the input volume. CNN's most important layer is the convolutional layer. It is being used to retrieve the characteristics of the input image or the upper layer's feature
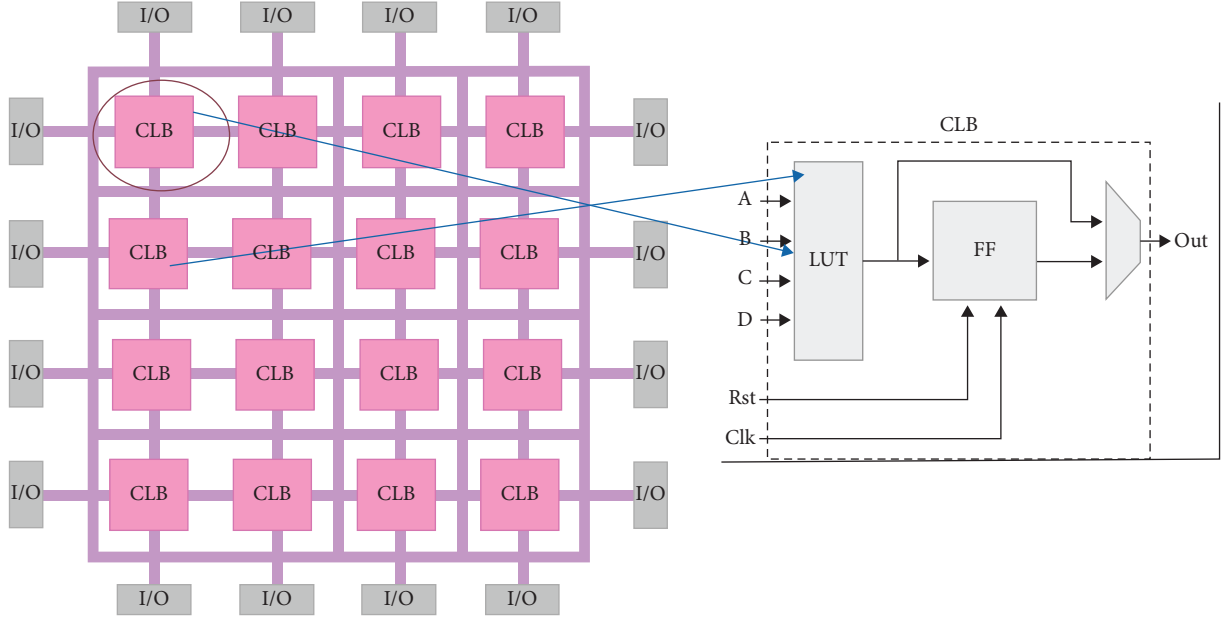
FIGURE 1: Overview of FPGA architecture, taken from [16].

map data [19]. The procedure is a three-dimensional convolution calculation based on input data and a huge variety convolution kernels, as well as the convolution operation is essentially a three dimensional multiply accumulate operation that could be described mathematically.

$$y_{out}(f_o, y, x) = \sum_{i_y=0}^{i-1} \sum_{i_x=0}^{i-1} w_l(f_o, i_y, i_x) \\ \times y_i(f_i, y + i_y, x + i_x) + b_i,$$

(1)

in which $y_i(f_i, y, x)$ as well as $y_{out}(f_o, y, x)$ refers neurons as input extracted feature f_i but also extracted feature f_o, respectively. $W_l(f_o, f_i, y, x)$ demonstrates the weights in the lth layer which is combined with f_i, as well as b_i would be a bias. The convolution filters are $i \times i$ in length.

2.3.2. Rectified Linear Unit Layer.

A recently proposed activation function in CNN is the Rectified Linear Unit (ReLU) that can be applied by thresholding a matrix at zero which is known to converge faster in training and has smaller computational complexity while the Sigmoid or tanh(x) activation functions involve expensive arithmetic operations [19]. The ReLU has become very popular in the last few years in convolutional neural network architecture. The equation of ReLU is very simple as follows:

$$f(x) = \max(0, x).$$

(2)

2.3.3. Pooling Layer.

As shown in Figure 2, the pooling layer is known as the down sample layer; it reduces extracted feature redundancy as well a network computational cost by minimizing extracted feature dimensions but rather effectively prevents overfitting. Pooling is among the common operators inside a convolutional neural network. Convolved
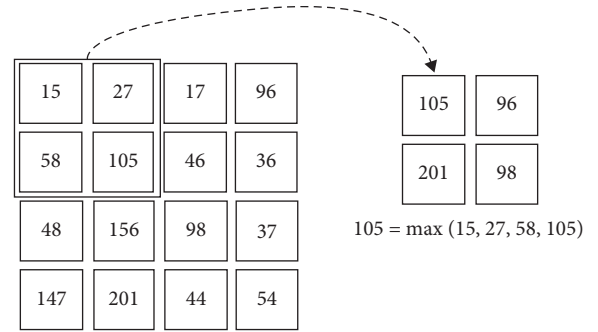
FIGURE 2: Max pool, taken form [19].

extracted features are compressed in a pooling layer by a dataset obtained near the area feature values [20]. Because images have the regional property, this operation is possible. The spatial size of feature values is reduced after the pooling operation, resulting in fewer computational tasks to perform in the flowing layer. The pooling operator's most common options include max pooling as well as average pooling. The term "max pooling" refers to the following:

$$O(i, j) = \max[O(i_o, j_o): i < i_o < i + p, j < j_o < j + p], \quad (3)$$

in which p is the operator's length and (i, j) is the vertical and horizontal index.

2.3.4. Fully Connected Layer.

The fully connected layer is the classical component of a feed-forward neural network, wherein every element inside the max pooling is linked to every component in the output nodes. The extracted features of the convolutional as well as max pooling require the input image's distributed high-level attributes. The FC layers were designed to combine such extracted features in order to

categorize the input into several classes. The forward throw of the *lth* FC layer is calculated as follows:

$$O^{l+1} = f\left(b^l + f^l + w^l\right), \tag{4}$$

where O^{l+1} is the output at $l + 1$ layers, f is the activation function, b^l is the bias for l^{th} layers, f^l is the feature map in l th layers, and W^l is the weights at the lth layers.

By adjusting the filter size of the convolution controller, an FC layer could be easily translated to the convolutional layer, which would be especially useful in practice.

2.3.5. Backpropagation. Back propagation has performed two updates that are for the weights and the deltas [21]. We are looking to compute $\partial E/\partial w_{m,n}^l$ which can be translated as the measurement of how a single pixel alters $w_{m,n}$ in the weight affects the loss function E. During forward propagation, the convolution operation ensures that the pixel $w_{m,n}$ in the weight, between an element of the weight and the input feature map element that it overlaps; a contribution is made in all the products [20]. Convolution between the input feature map of dimension $H \times W$ and the weight of dimension $k_1 \times k_2$ produces an output feature map of size $(H - k_1 + 1)$ by $(W - k_2 + 1)$. By applying the chain rule in the following way, the gradient component for the individual weights can be obtained [9].

$$\frac{\partial E}{\partial w_{m,n}^l} = \sum_{i=0}^{H-k_1} \sum_{j=0}^{W-k_2} \delta_{i,j}^l \frac{\partial x_{i,j}^l}{\partial w_{m,n}^l}. \tag{5}$$

The summations represents a collection of all the gradients $\partial_{i,j}^l$ coming from all the outputs in layer *l*.

3. Literature Review

Recent FPGAs had also provided a significant design space for a convolutional neural network due to an increase throughout FPGA fabric density as well as reducing transistor scale. The work by Tapiador et al. [6] implemented a depth-wise separable convolution with a high rate of resources and also significantly increases bandwidth as well as accomplishes a complete pipeline through parameter tuning and through a streaming data interface and the on ping-pong. In the work by Kaiyoua et al. [22], CNN models and CNN-based implementations have been distinguished. The requirements for memory, computation, and system reliability for mapping CNN on embedded FPGAs were summarized. Requirement analysis: they proposed Angel-Eye, which is a programmable as well as configurable CNN hardware accelerator combined with quantization method, compilation tool, as well as a data quantization technique. The compilation tool converts a specific CNN model into the hardware configuration. They were tested on the Zynq XC7Z045 platform and outperformed; peer FPGAs on same network have same of performance as well as power efficiency by 6*x* and 5*x*, respectively. In the work by Naveen Suda et al. [5], FPGA throughput is optimized on large-scale CNN with 3D convolution as matrix-multiplication. Their work demonstrated that ImageNet classification on the P395-D8

board can achieve a peak performance of 136.5 GOPS for convolution operations and 117.8 GOPS for the entire VGG network.

In terms of the processing time, the FPGA implementation has almost the same performance as the GPU implementations although the FPGA's memory bandwidth is much smaller and has much high energy efficiency than the GPU's one. FPGAs will be advantageous in the high-performance computing scope for these reasons because they provide reprogrammable hardware as well as low power consumption, and FPGA implementation is a cost effective also fast [12] while OpenCL enhances the code portable as well as programmable of FPGA, which greatly reduces the time and complexity programming process and it comes at the best of performance [8].

4. Methodology

In this section, we would go over the architecture in general, including convolution, input max pooling, and backward and output kernels.

4.1. Accelerator Design. The overall system design flow as well as both host and device system section of the OpenCL kernels is created with the Intel FPGA SDK for the OpenCL enhanced version channel. The hardware accelerators design has five kernels, such as forward convolution, backward convolution, pooling, input, and output. The input and output kernels have been used to transfer extracted features as well as weight from and to the main memory, which brings some kernels with high-throughput sequencing data. The convolution kernel is designed to speed up the most parallelize computations in CNNs, which typically include the convolution operation and the FC layer [7].

The Max-pool part works to dwindle the dimension of the information by combining the outputs of neurons into a single within the another layer and undersampling operations specifically on the yield data stream of the convolutional part. The cascaded kernels shape a channel, which can operate the essential CNN operations without the requiremet of putting away interlayer information backmost to external memory. So, every convolution channel has a computing unit, and the kernel is made up of many computing units to do parallel convolution [9]. Both of input and output kernels are a most vital kernel which are utilized for a data movement and a kernel that is outlined to bring or store information from or to a main memory for the computing path. The input kernels begin with such a global work items in convolution configuration whereas the output kernel is operating in an NDRange unveiling with global work items. To enable concurrent work group processing, the work items have been organized up into multiple running in parallel work groups, also with a local work group length of (i, i). The convolution filters size is 3×3, which minimizes computational costs and weight sharing that to lower back-propagation weights. The number of pixels shifted over the input matrix is referred to as the stride, and we use the stride size as 2×2 to modify the amount of movement over the image.

4.2. Forward Convolution Kernel. The forward convolution kernel performs a convolution operation. The forward convolution kernel performs a convolution operation. At each position, the multiplication between each element of the kernel and the input feature map element is computed

and the results are summed up to obtain the output at that current location. The convolution operation is essentially a three-dimensional multiply accumulate (MAC) operation, which can be defined as

$$y_o(fe_o, y, x) = \sum_{fe_i=1}^{C_l} \sum_{i_y=0}^{i-1} \sum_{i_x=0}^{i-1} w_l(fe_O, i_y, i_x) \times y_i(fe_i, y + i_y, x + i_x) + b_i, \tag{6}$$

in which $y_i(fe_i, y, x)$ as well as $y_o(fe_o, y, x)$ refers neurons as input extracted feature f_i but also extracted feature fe_o, respectively. $W_l(fe_o, fe_i, y, x)$ demonstrates the weights in the lth layer which is combined with fe_i, as well as b_i would be a bias.

4.3. Input Kernel. The algorithm 1 shows that the input kernel is used for reading input extracted features and relates filters from memory, along with feeding weight into the local buffer and obtaining extracted features and caching them in the local buffer. Because an input extracted feature is recycled by numerous different filters, the input array is cached in local memory for access during data processing and to reduce the access of global memory.

(1) Get global and local index of work item

(2) Calculate location for input features and filters using index

(3) Bring input features into the local memory

(4) Bring filter into the local memory

(5) #progma unroll

(6) for each component i in both input feature and filter do

(7) Load weight into weight buffer

(8) Fetch the weight and bias by fetcher

(9) end for.

4.4. Pooling Kernel. The Pooling part performs to reduce the dimension of the weight by combining the outputs of neurons into a single within the another layer and undersampling operations specifically on the output data stream of the convolution kernel. The pooling layer reduces the convolutional outcomes while using the average or maximum value of elements in an area that is dependent on subsequent iterations. A shift registers with the depth that is developed for caching the accumulating data, similar to such convolutional layer. Then, depending on the pooling method, accumulating operations are performed on the shift register.

4.5. Output Kernel. The output kernel reads backproagation results from the accumulation channel and writes them back to global memory and then outputs to a local buffer, then extracts the data from the buffer, and copies it back to DDR.

This work makes use of batch processing to reduce the time it takes for filters to be reused in FC layers. As a result, in the FC layer output kernel, N batch sets of results must be collected and written. It processes one set of results for the additional layer. The kernel is constructed in an NDRange manner, executing with work items in parallel, so the output processing is entirely independent.

4.6. Backward Convolution Kernel. This kernel reads the result from the max pooling buffer channel as well as performs two functions: error δ calculation and partial derivatives ΔW and ΔE calculation, both of which are cross-correlation processes. The cross-correlation operation can be implemented by reversing the data in the convolution kernel. The difference in resource usage is that while calculating the derivatives, we require two input buffers for both δ^l and δ^{l-1}. Convolution between the input feature map of dimension $H \times W$ and the weight of dimension $k_1 \times k_2$ produces an output feature map of size $(H - k_1 + 1)$ by $(W - k_2 + 1)$. By applying the chain rule in the following way, the gradient component for the individual weights can be obtained [9].

$$\frac{\partial E}{\partial w_{m,n}^l} = \sum_{i=0}^{H-k_1} \sum_{j=0}^{W-k_2} \delta_{i,j}^l \frac{\partial x_{i,j}^l}{\partial w_{m,n}^l}. \tag{7}$$

5. Optimizations for Performance

In this section, we will discuss performance optimization techniques such as throughput maximizing, quantization, memory communication, parallelism in convolution neural networks, and converting fully connected layer to convolution layer.

5.1. Throughput. To keep moving forward, the accelerator's throughput, SIMD, and concurrent computing units are announced. The input kernel retrieves the SIMD and sends it to numerous computing units in the convolution. By adjusting the value of the SIMD as well as the number of computing units that is deployed, design could obtain scalable performance and hardware costs without requiring changes to the kernel code.

5.1.1. Computing Unit. The FPGA chip's resources are limited. If hardware resources are required for the optimization techniques, each kernel could have multiple compute units generated. This necessitates the creation of multiple copies of the various transmission lines. Even so, multiple computing units could not always improve throughput linearly since all computing units communicate over the global memory bandwidth. This causes memory access contention among computing units.

5.1.2. Single Instruction Multiple Data (SIMD). To increase the data processing performance of an OpenCL kernel by processing various work items can be accessed by a single instruction multiple data (SIMD) approach without annually vectorizing the kernel code. The largest amount of work items per workgroup that the Intel FPGA SDK for OpenCL compiler could execute SIMD or vectorized was determined. The work group size that could be used is defined by the compiler, and the local work size argument is used to clEnqueueNDRangeKernel. The workgroup length can be allowed to pass to clEnqueueNDRangeKernel as such local work length argument. The above enables the compiler to adequately enhance the generated kernel code.

5.1.3. Loop Unrolling. The several loop iterations in the device code could have an impact on the kernel performance. The loop unrolling method could assign the most hardware resources and minimize or even eliminate the loop queue, that is, increase the throughput in a linear manner. This approach supports memory coalescing as well that also reduces memory transaction cost.

5.2. Quantization Technique. In general, artificial neural deployments, including convolutional neural networks, make use of a 32-bit floating point. The circumstance, even so, has been transformed. Several more latest FPGA works on convolutional neural networks had also centered to use the fixed-point representation of the extremely narrow bit width, which now has accuracy reduction [23–25]. However, nevertheless, low-bit reduction-based designs demonstrate exceptional performance and energy efficiency; this indicates that extremely low-bit width is an useful solution for higher efficiency design [23, 26, 27].

[IL.FL], from which IL seems to be the total number of integer bits and FL has been the total number of fractional bits would be a fixed-point number structure. The overall number of bits is calculated as the sum of IL and FL as well as the fixed-point number has an exactness of 2^{-FL} and the scope could be described this way: $[-2^{IL-1}$ and $2^{IL-1} - 2^{-FL}]$ [23]. The fixed point is the hardware-friendly as well as enables so much logic resources on FPGAs, allowing for increased parallel computing [28]. This even decreases the chip's memory usage and bandwidth needs. Even so, as in fixed-point deployment, we would use a fixed-point which was with static configuration to create cost effective and much more precise hardware kernels [15]. In overall,

quantization is the most significant element in accelerating huge CNNs on the FPGA platform.

5.3. Memory Communication. Because several developments are limited by memory bandwidth, the other option is to use efficient memory access to reduce communication cost. Several developments are limited by memory bandwidth, the other option is to use efficient memory access to reduce communication costs.

5.3.1. Memory Alignment. Here, on host side, memory allocated would have to be at least 64-byte aligned. This significantly improves the transmission efficiency of DMA transmitting on the host-FPGA communication. The allocation can be executed in Linux using the POSIX mem-align function, which is supported by GCC, or Windows use that aligned malloc function, which is held by Microsoft.

5.3.2. The Local Memory Caching. Global memory, constant memory, local memory, and private memory are the four areas of the OpenCL memory model. Local memory, which would be executed in the on-chip Random access memory block, does have significantly decreased latency and high bandwidth than main memory. As a result, we can cache global memory which requires multiple accesses previous to computation using local memory. Those certain cached local memories have been viewable to everyone, work items in the same workgroup when data parallelism is enabled. By minimizing the memory access, the use of local memory would improve kernel performance.

5.4. Parallelism in Convolutional Neural Network. Those processing, which would include reading, convolving, pooling, and writing back, are data independent of varying extracted features. As a result, the entire output extracted feature can be vectorized along the N dimension, with each section processed on a different data path. This can be executed by a computing unit in OpenCL that would significantly enhance the proposed design throughput [10]. Furthermore, every convolution operation of an extracted feature unit consists of stage element-wise multiplication of input extracted feature and filters, followed by the accumulation of the product of these operations. [19]. In the first process, multiplication is completely independent and could be performed using a data parallelism technique.

5.5. Changing FC Layers to Convolution Layers. Fully connected layers and convolution layers have the same working order form, which entailed multiplying and adding. It should be noted that because the only difference between the fully connected and the convolution layer would be that the neurons in the convolution layer are only connected to a local region at the input and that many of the neurons in the convolution layer volume share parameters. The fully connected (FC) layer operates on a flattened input, with each input connected to all neurons. Dot products, on the other

hand, are always computed by neurons in both layers, so their functional form is similar. There are two approaches for changing FC layers to convolution layers. First, choose a convolution layer kernel filter with the same length also as input feature's map, and secondly, by using 1×1 convolutions with multiple channels.

6. Experimental Setup

The Terasic DE1 SoC Development Kit (DK) of FPGA board is used to implement the experiments. DE1 SoC would be a powerful hardware design platform based on Intel System-On-Chip (SoC) FPGA. The DE1 SoC board uses several features which enable designer to complete a broad range of designing circuits projects.

The terasic DE1 SoC board has M10K-10-kbit memory blocks including soft error correction code (ECC), as well as a 400 MHz/800 Mbps interface of an external memory and 64 MB of the SDRAM, 1 GB ($2 \times 256M \times 16$) of DDR3, and micro SD card port on Hard Processor System (HPS) memory [29]. The Intel cyclone V SoC 5CSEMA5F31C6 has 85K programmable logic elements, 4,450 Kbits of memory embedded, 6 fractional phase locked loops (PLLs), dual-core ARM Cortex-A9 (HPS), and 2 memory controllers based on TSMC's 28-nm low power (28LP) process technology. The architecture of a DE1 SoC includes two USB 2.0 Host ports (ULPI interface with USB type A connector) [29]. As communication ports, connectors, displays, switches, buttons, indicators, audio, and video inputs, G-Sensor on HPS and UART to USB (USB Mini-B connector), 10/100/1000 Ethernet, PS/2 mouse/keyboard, IR emitter/receiver, and I2C multiplexer are used. The accelerator boards communicate with the host through the use of an 8-lane PCI express link.

We use Intel SDK for OpenCL intelFPGA_Standard_18.1.0 build 625. The Intel FPGA SDK for OpenCL Standard Version includes programs, drivers, development kit library resources as well as files, and much more. The Intel SDK for OpenCL Standard_18.1.0 has logic components such as offline Compiler translates, a set of commands, host runtime providing the OpenCL host, and runtime API for the OpenCL host code. We used the Board Support Package (BSP) 18.1 version for de1soc board BSP from Terasic and Intel SDK for OpenCL the intelFPGA_Standard_18.1.0 with 625 buildings is used. Additionally, (Intel® CoreTM i5-4300) CPU and (AMD Radeon (TM) R5 M330) GPU are used.

7. Result and Discussion

In this section, we evaluate the performance of our proposed system with the different design specifications. The objective of this exercise is to learn the resource utilization and performance figures for combinations of design specifications. We employ two well-known CNN models for the possible combinations of convolutional neural network design specifications.

7.1. Design Performance and Analysis. We advanced to evaluate the accuracy of our design on the ImageNet ILSVRC-2012 data set, where it contains up to 1.2 million training and 50k validation instances. An AlexNet Caffe model, that has 61 million parameters as well as a top-1 accuracy of 57.2% and a maximum classification of top-5 accuracy of 80.3%, has been used as a reference model. On the same ILSVRC-2012 data set, we furthermore examined a larger, more latest network, VGG-16 [30]. The VGG-16 does have 138 million parameters and much more convolutional layers, but still only three fully connected layers at the moment [12]. The accuracy of our work was assessed with executing our models on 728K training and 50K validation samples from the ImageNet 2012 data set. The accuracy comparision for AlexNet model in Figure 3 and the accuracy comparision for VGG-16 model demonstrate the accuracy of various quantization compression rates.

Figure 3 and Figure 4 demonstrate the accuracy of various quantization compression rates. As shown, the model's accuracy starts to decline considerably while compressing below 8-bit data quantization of its base accuracy. The difference between the Caffe tool using AlexNet model with 32 bit floating point and the 32 bit floating point FPGA design on top-1 and top-5 accuracies is 0.5% and 0.59%, respectively. The difference between a 16-bit fixed point Caffe tool and FPGA design on top-1 accuracy is 0.59% and top-5 accuracy is 0.9% accuracy loss compared to the reference design. The accuracy difference between 8 bit Caffe and FPGGA implementation design on top-1 and top-5 accuracies is 0.77% and 0.5%, respectively. The accuracy difference between 4 bit Caffe tool and FPGGA implementation design on top-1 and top-5 accuracies is 2.05% and 1.19%, respectively. Therefore, the accuracy of our implementation is excellent. As a result, the exactness of our implementation is comparable to baseline.

7.2. Computation Throughput and Energy Efficiency. In this subsection, we will discuss the computation throughput as well as the energy efficiency of the system. Figures 5 and 6 show throughput, and Figures 7 and 8 depict energy efficiency.

7.3. Computation Throughput. In Figures 5 and 6, our experiments have show that with low-bit width quantization, we can achieve a high throughput in results. The low-bit width quantization techniques have significant benefits because it allows for high memory cache to be used as well as removes memory constraints in deep learning methods. This enables faster data movements and more efficient computation of the throughput in hardware acceleration. And it enables the device to do more operations per second, significantly speeding up workloads. Because of these advantages, low-bit width implementations are likely to become common in training and inference, particularly for convolutional neural networks.
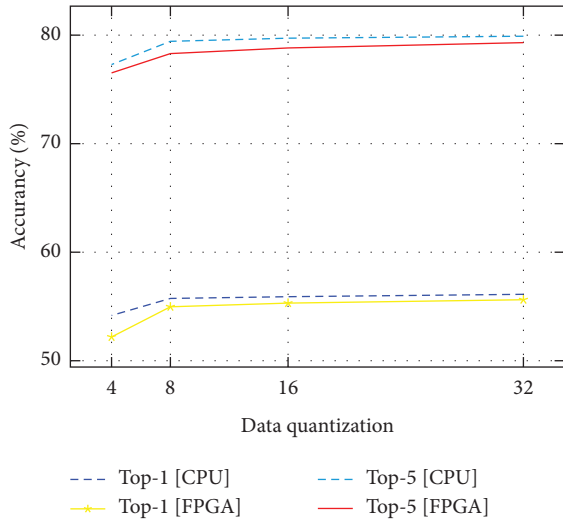
FIGURE 3: The accuracy comparison for AlexNet model.
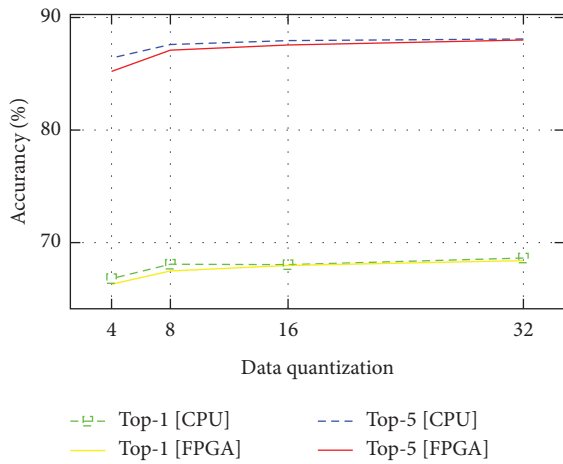
FIGURE 4: The accuracy comparison for VGG-16 model.

FIGURE 5: Throughout with difference data quantization with Caffe [CPU].
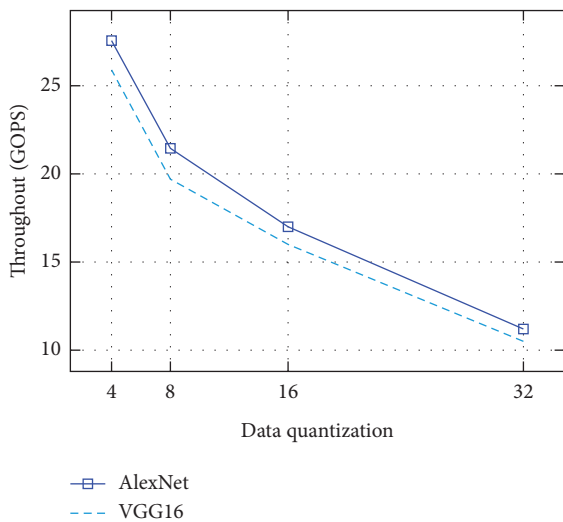
FIGURE 6: Throughput with difference data quantization on FPGA.
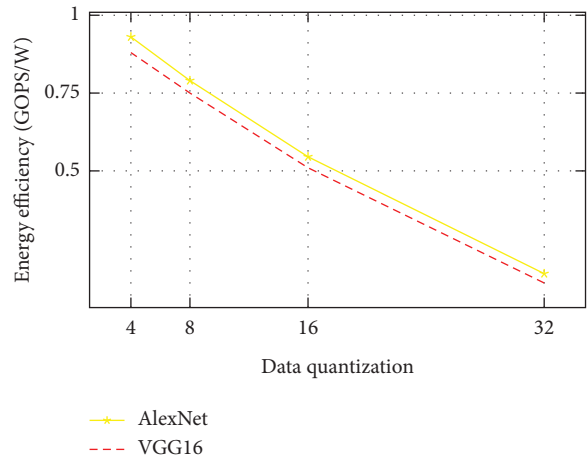
FIGURE 7: Energy efficiency with difference data quantization with Caffe [CPU].
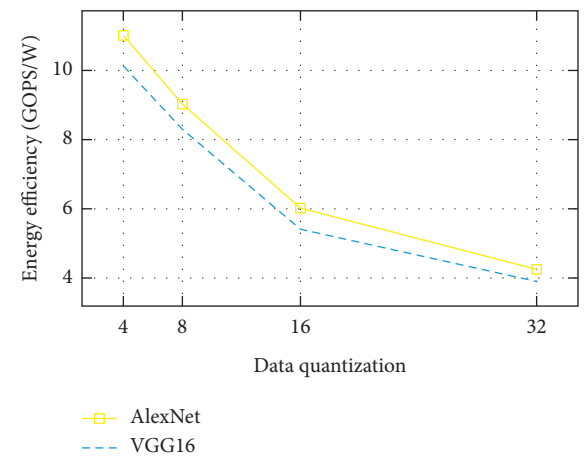
FIGURE 8: Energy efficiency with difference data quantization on FPGA.

7.4. Computation Efficiency. Figures 7 and 8 demonstrate that the low-bit width quantizataion neural networks improve power efficiency. As we have discussed in the subsection of computation throughput, it reduces memory access costs by enabling high memory cache usage and increasing compute efficiency. Using low-bit quantization can reduce power consumption and save significant energy. Low-bit width quantization uses less energy and enhances compute efficiency, resulting in lower power consumption. Furthermore, decreasing the number of bits used to represent the neural network's parameters results in less memory storage.

Generally, among all the data quantization as observed from Figures 5 to 8, low-bit width-based designs demonstrate exceptionally good speed and energy efficiency. This indicates which extremely low-bit width is a likely answer for high performance. However, the extremely low-bit width has accuracy reduction.

7.5. Resource Utilization. Table 1 shows the trained CNN on FPGA resource usage. During training, CNN on FPGA consumes huge computational resources. We trained our models on the de1_SoC board before changing any parameters, and the resource usage is illustrated in Table 1.

Our design, as stated in section III, includes two major variables: the number of computing units and the number of SIMD. The replication of full data paths is the total number of computing units, that also controls the balance among both resource usage. Additionally, a compute unit may be composed one or more processing elements depending on our design choice. Having more processing elements per compute unit can significantly raise the data processing speed by allowing single instruction multiple data (SIMD) execution mode.

The number of SIMD processing elements is a design choice that also allows for contiguous memory retrieval which can enhance memory utilization efficiency. In the design, we were using a static configuration number of SIMD units, which allowed us to work with restricted on-board resources. We investigate how well the number of computing units and SIMD impact the De1 Soc-based board's resource utilization and throughput.

In order to achieve the maximum performance of our design, we configured the SMID as fixed as well as varied the number of computing units. When both parameters grow, there is also a growth in resource usage. Furthermore, with data path replication in the framework, the number of computing units does have a greater effect on resource utilization than that of the number of units. Whenever these variables are increased, it is simple to see even a linear improved performance in throughput. However, because of the limited resources on the DE1 SoC, the integration of computing is equal to sixteen as well as SMID sixteen results in successful synthesis both on fixed point and floating point.

7.6. Power Measurement. The power consumption is an important element in hardware accelerator performance. The power drain on one of the devices tells us how hard it is working and how power-intensive the design would be. This is especially essential for evaluating deep learning applications for hardware accelerators, where power consumption is a major consideration. We measure performance and power consumption by using the Perf performance analysis tool for Linux. The idle CPU-only system absorbs 50.70 W before the FPGA accelerator board is installed on the system. When using Caffe tools to run AlexNet and VGG-16 models, the average power utilization starts to rise to 109.2 W. Whenever a DE1 SoC-based FPGA board is properly configured, the idle power consumption rises to 63.40 W. Throughout CNN kernel implementation, the overall power usage of the hardware acceleration rises to 78.2 W by averages. Thereby, a power use for running a CNN framework on the a DE1 SoC-based board is (78.20–50.70) = 27.5 W.

7.7. Comparative Discussion on Previous Work and Other HPC Platform Design. In this section, we would first compare our implementations with previous FPGA research. The following is a comparison with similar designs focused on other high-performance computing platforms, such as CPUs and also GPUs.

7.7.1. FPGA-Based Design. We contrast the proposed models' efficiency to that of a number of other recent FPGA-based CNN design features. To determine the throughput, divide the total floating point numbers or fixed-point operations through the entire execution time and then use GOP/S as a unit for floating point as well as fixed-point operations in our implementation design. Zhang et al. [31] implemented a convolution layer which obtained 61.62 GOPS again for single precision floating point design. Similarly, the work by Yufei Ma et al. [13] reported 134.1 GOPS and 117.3 GOPS on a convolution layer for AlexNet and NIN model, respectively, while they achieved the overall performance 114.5 GOPS and 117.1 GOPS for AlexNet and NiN model, respectively. Our throughput from the 4-bit fixed point on DE1_SoC for AlexNet model 203.75 GOPS and also for VGG-16 model 196.50 GOPS on DE1_SoC. Our work gained 1.78x more throughput over the work by Naveen Suda et al. [5] with only using 85 DSP blocks. Furthermore, our design outperforms the RTL design in [13] by 1.51x on the different boards, demonstrating that OpenCL-based designs can compete with RTL designs. When compared to other designs, our DE1 SoC design has had the highest throughput, and there is still room for improvement.

7.7.2. Other HPC Platform-Based Design. We as well introduce energy consumption as both a measure for evaluation, which would be the ratio of throughput to power consumption (GOPS/Watt). In terms of throughput, the GPU is the best alternative, followed by one FPGA design, as shown in Table 2. Power usage, on the other hand, is an important measure to take into account in modern digital design. The GPU absorbs 3.709X so much energy than that of the FPGA, and the FPGA is 22.613x more efficient than the CPU.

TABLE 1: Resource usage CNN training.

Data quantization (bit)	ALUTs	DSP	FFS	M10 K
32	139913	85	172677	497
16	108313	76	144798	422
8	88712	64	126918	346
4	74511	52	91158	187

TABLE 2: Compare with other devices.

Platform	CPU Intel® Core i5-4300	GPU AMD Radeon (TM) R5 M330	FPGA DE1_SoC
Technology	22 nm	28 nm	28 nm
Power (Watt)	58.5	94.50	18.5
Throughput (GOPS)	28.50	280.60	203.75
Energy efficiency (GOPS/W)	0.487	2.969	11.013

8. Conclusion

In this work, we show a training and classification of a deep neural network that use the Intel, FPGA OpenCL SDK. To determine the best design requirements to speed up the CNN model for training while using constrained FPGA resources, we proposed a design space exploration methodology for energy efficiencies and resource utilization. We implemented CNN models such as AlexNet and VGG on the DE1 SoC FPGA board using the proposed approach as well as gained higher performance when compared to earlier work. As we compared with the other platform, the CNN training on FPGA consumes less power consumption and training time. Our findings indicated that FPGAs could obtain greater power or energy efficiency than GPUs, which typically restrict improvement only to power efficiency. We noted that it is mainly due to the huge difference in maximum compute performance as well as the external memory bandwidth between FPGAs and GPUs.

Generally, our designs achieve 203.75 GOPS on Terasic DE1 SoC with the AlexNet model and 196.50 GOPS with the VGG-16 model on Terasic DE-SoC. This, as far as we know, outperforms existing FPGA-based accelerators. Compared to the CPU and GPU, our design is 22.613X and 3.709X more energy efficient, respectively.

Data Availability

The source of the author's framework along with the datasets and analysis during the current study is already publicly available on https://image-net.org/challenges/LSVRC/2012/index php which is maintained by Princeton University and Stanford University. Quartus-18.1.0.625 software was used for processing and classification purposes during the author's research experiment.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[3] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna,, "Rethinking the inception architecture for computer vision," in *Proceedings of the Ieee Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015.

[4] J.-H. Lin, T. Xing, R. Zhao et al., "Binarized convolutional neural networks with separable filters for efficient hardware acceleration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 27–35, Honolulu, HI, USA, July 2017.

[5] N. Suda, V. Chandra, G. Dasika et al., "Throughput-optimized opencl-based fpga accelerator for large-scale convolutional neural networks," in *Proceedings of the 2016 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pp. 16–25, Monterey CA, USA, February 2016.

[6] R. Tapiador, A. Rios-Navarro, A. Linares-Barranco, M. Kim, D. Kadetotad, and J.-s. Seo, "Comprehensive evaluation of opencl-based convolutional neural network accelerators in xilinx and altera fpgas," 2016, https://arxiv.org/abs/1609.09296.

[7] W. Ding, Z. Huang, Z. Huang, L. Tian, H. Wang, and S. Feng, "Designing efficient accelerator of depthwise separable convolutional neural network on fpga," *Journal of Systems Architecture*, vol. 97, pp. 278–286, 2019.

[8] J. Zhang and J. Li, "Improving the performance of opencl-based fpga accelerator for convolutional neural network," in *Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pp. 25–34, Monterey CA USA, February 2017.

[9] M. Mathieu, M. Henaff, and Y. LeCun, "Fast training of convolutional networks through ffts," 2013, https://arxiv.org/abs/1312.5851.

[10] U. Aydonat, S. O'Connell, D. Capalija, A. C. Ling, and G. R. Chiu, "An opencl deep learning accelerator on arria 10," in *Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, pp. 55–64, Monterey, CA, USA, February 2017.

[11] S. Tatsumi, M. Hariyama, M. Miura, K. Ito, and T. Aoki, "Opencl-based design of an fpga accelerator for phase-based correspondence matching," in *Proceedings of the International Conference On Parallel And Distributed Processing Techniques And Applications (PDPTA). The Steering Committee of the World Congress in Computer Science*, p. 90, Las Vegas, USA, July 2015.

[12] H. Li, *Acceleration of Deep Learning on Fpga*, University of Windsor, Canada, 2017.

[13] Y. Ma, N. Suda, Y. Cao, S. Vrudhula, and J.-s. Seo, "Alamo: fpga acceleration of deep learning algorithms with a modularized rtl compiler," *Integration*, vol. 62, pp. 14–23, 2018.

[14] K. Guo, S. Zeng, J. Yu, Y. Wang, and H. Yang, "[dl] a survey of fpga-based neural network inference accelerators," *ACM Transactions on Reconfigurable Technology and Systems*, vol. 12, no. 1, pp. 1–26, 2019.

[15] A. Shawahna, S. M. Sait, and A. El-Maleh, "Fpga-based accelerators of deep learning networks for learning and classification: a review," *IEEE Access*, vol. 7, pp. 7823–7859, 2019.

[16] W. Stallings, *Computer Organization and Architecture: Designing for Performance*, Pearson Education India, Noida, 2003.

[17] K. Morris, "The Path to Acceleration: Altera Bets on Opencl," 2012, http://www.eejournal.com/archives/articles/20121106-opencl/.

[18] L. Howes and A. Munshi, "The opencl specfification," 2013, https://www.khronos.org/registry/Opencl/Specs/Opencl-1.2.pdf.

[19] B. Liu, D. Zou, L. Feng, S. Feng, P. Fu, and J. Li, "An fpga-based cnn accelerator integrating depthwise separable convolution," *Electronics*, vol. 8, no. 3, p. 281, 2019.

[20] Z. Liu, Y. Dou, J. Jiang, Q. Wang, and P. Chow, "An fpga-based processor for training convolutional neural networks," in *Proceedings of the 2017 International Conference on Field Programmable Technology (ICFPT)*, pp. 207–210, IEEE, Manhattan, NY, USA, December 2017.

[21] W. Zhao, H. Fu, W. Luk et al., "F-cnn: An fpga-based framework for training convolutional neural networks," in *Proceedings of the 2016 IEEE 27 th International Conference on Application-specific Systems, Architectures and Processors (ASAP)*, pp. 107–114, IEEE, Manhattan, NY, USA, July 2016.

[22] K. Guo, L. Sui, J. Qiu et al., "Angel-eye: a complete design flow for mapping cnn onto embedded fpga," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 1, pp. 35–47, 2018.

[23] L. Jiao, C. Luo, W. Cao, X. Zhou, and L. Wang, "Accelerating low bit-width convolutional neural networks with embedded fpga," in *Proceedings of the 2017 27th International Conference on Field Programmable Logic and Applications (FPL)*, pp. 1–4, IEEE, Manhattan, NY, USA, September 2017.

[24] S. Han, H. Mao, and W. J. Dally, "Deep compression: compressing deep neural networks with pruning, trained quantization and huffman coding," 2015, https://arxiv.org/abs/1510.00149.

[25] A. Prost-Boucle, A. Bourge, F. Pétrot, H. Alemdar, N. Caldwell, and V. Leroy, "Scalable high-performance architecture for convolutional ternary neural networks on fpga," in *Proceedings of the 2017 27th International Conference on Field Programmable Logic and Applications (FPL)*, pp. 1–7, IEEE, Manhattan, NY, USA, September 2017.

[26] H. Nakahara, H. Yonekawa, H. Iwamoto, and M. Motomura, "A batch normalization free binarized convolutional deep neural network on an fpga," in *Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, p. 290, Association for Computing Machinery, New York, NY, USA, February 2017.

[27] H. Nakahara, T. Fujii, and S. Sato, "A fully connected layer elimination for a binarizec convolutional neural network on an fpga," in *Proceedings of the 27th International Conference on Field Programmable Logic and Applications (FPL)*, pp. 1–4, IEEE, Ghent, Belgium, September 2017.

[28] D. Wang, J. An, and K. Xu, "Pipecnn: an opencl-based fpga accelerator for large-scale convolution neuron networks," 2016, https://arxiv.org/abs/1611.02450.

[29] "Terasic de1soc user manual," 2014, http://www.ee.ic.ac.uk/pcheung/teaching/ee2_digital/de1-soc_user_manual.pdf.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[31] M. Zhang, L. Li, H. Wang, Y. Liu, H. Qin, and W. Zhao, "Optimized compression for implementing convolutional neural networks on fpga," *Electronics*, vol. 8, no. 3, p. 295, 2019.

Research Article

IoT Speech Recognition Application in Mass Sports Data Monitoring Based on Dynamic Adaptive Recommendation Algorithm

Meng Song [ID] [1,2]

[1]*Tianjin University of Sport, Tianjin, China*
[2]*Inner Mongolia University of Science and Technology, Inner Mongolia, China*

Correspondence should be addressed to Meng Song; songmeng09@imust.edu.cn

In this paper, the existing dynamic adaptive recommendation methods are studied, which combine the practical application scheme of transforming the actual dynamic adaptive recommendation problem into user microblog information. After that, a dynamic adaptive weight fusion method is proposed and based on experimental verification, a real-time dynamic adaptive recommendation system is finally designed. The speech recognition of the Internet of Things takes natural language problems as the research object for a long time and takes the sound signal as the research topic. This paper analyzes the application of dynamic adaptive recommendation and Internet of Things speech recognition in mass sports data monitoring. The simulation results show that the system in this paper is convenient for users to monitor the exercise indicators in real time through the mobile client, and at the same time query the exercise historical data and compare the exercise data through the network terminal, thereby improving the exercise method and exercise load. Users can access the motion monitoring module and see the past floating state of motion parameters more intuitively than graphs, contains queries for metrics such as heart rate, body temperature, kinetic energy, pulse, and weight. Due to the diversity and complexity of people's differences, personal characteristics and business environments, sports data monitoring systems also need to be designed according to the scope of use. This paper analyzes the requirements for a motion data monitoring system and provides the system architecture design and basic data for producing detailed information for the system.

1. Introduction

This paper mainly proposes a dynamic adaptive recommendation method that meets the user's interest, content, and site operator's authority [1]. Based on the rearrangement of the adaptive recommendation method, the characteristics of the dynamic adaptive recommendation method are integrated [2]. In order to prove its practicability, it was verified by coding. After proving the effect of the method, this paper designs and develops a personalized prototype system. The experimental environment of this article is (1) hardware configuration: Intel(R) I5-3110M3.2 GHz, 12G memory; (2) operating system: Win764 bit; (3) development environment: Eclipse, MySQL, Navicat, editolus R2012. At this stage, most of the Internet of Things speech recognition

technologies are based on statistical models. From the principle of speech generation, speech recognition can be roughly divided into two parts: speech layer and language layer [3]. The difficulty of current Internet speech recognition technology can be solved from the aspect of speech characteristics. Audio model is the most basic model that represents audio information, and it is also an important part of the system [4]. The type of training unit directly depends on the method of feature extraction. Feature extraction requires the extraction function not only to effectively reduce the data volume of the digital audio signal, but also to ensure the validity of all the original audio information [5]. However, due to the serious deficiencies in the development system of mass sports at this stage, the specific manifestations are incomplete basic facilities, imperfect

organizational structures, insufficient funds, lack of management talents, and lack of "work is sports" [6]. On this basis, we can take advantage of the favorable promotion of new construction to actively study the development path of mass sports in the field of construction, which has important significance in practical applications [7]. The sports data monitoring system server must provide real-time services for many sports users [8]. In order to realize real-time data monitoring, the system server provides stable and efficient features with low latency. At the same time, online operations are supported. By analyzing two traditional synchronization models, it is difficult for the traditional synchronization model to meet the scalability of complex network services [9]. The SEDA concurrency architecture model can flexibly adapt to system requirements and expansion, and combines the advantages of multithreading and event-driven that can maintain excellent performance in the environment. This will simplify the development of the system server, while also increasing the speed of the server, which contributes to the performance of the entire system [10]. Provides a strong foundation to optimize and promote server performance.

2. Related Work

The literature introduces the research content of the sports health evaluation model and explains the concept of sports health evaluation. Due to the complexity and randomness of the movement process, we can use the Markov statistical model for analysis. Next, the basic theory of the MARKP model is described, and the characteristics of the Markov model are analyzed [11]. Finally, for sports health assessment and various requirements of the group, an exercise model of health assessment is proposed. The literature introduces a real-time monitoring system for data collection hardware equipment, user equipment data, and sports data [12]. The data collection hardware equipment can be used to obtain real-time user data and display sports data. Among them, the data collection device uses 3D acceleration sensors to analyze the characteristics of people participating in the movement. These data are converted into a base station, and the base station sends specific data to the system communication service memory, combining the characteristics of the user's body and the energy consumption of the exercise, and finally sends the obtained data to the mobile monitoring terminal [13]. The literature introduces the operation time and energy consumption data obtained from the exercise data real-time monitoring system and proposes a health assessment model to explain the data consumption and personalized exercise to achieve the purpose of providing recommendations [14]. The literature analyzed the characteristics of energy consumption and individual height and weight during exercise, and proposed a health evaluation model for sports population [15]. The purpose is to eliminate the differences caused by individuals in sports, make an objective evaluation of the athletes, and give a reasonable evaluation effect to the sports team. The literature describes the basic principles of neural networks and the basic theories of audio signal processing [16]. First of all, based on the

principle of speech recognition, the speech recognition has been improved. Including digital sampling of audio signals and digital sound signal analysis, while spectrum analysis technology is suitable for signal content analysis and extraction. The literature introduces the back propagation network model and the deep belief network model, and gives some opinions and suggestions, many of which still need to be in-depth research and testing. In terms of speech processing, the samples used are collected in an ideal indoor environment. In practice, the samples may have obvious noise. Therefore, the influence of noise should be eliminated first in the experiment process to improve practicability.

3. Theoretical Basis of Dynamic Adaptive Recommendation and Internet of Things Speech Recognition

3.1. Dynamic Adaptive Recommendation. In this article, taking into account the degree of decline in the time value, it is divided into the following processes:

(1) Based on the original data, calculate sports popularity and time decay factors, and save them in their respective popularity databases and damping factor databases

(2) Use the neighborhood method to calculate the user motion similarity matrix and store it in the neighborhood method similarity database

(3) The loss function is defined by the matrix factorization method and stored in the loss function library of the matrix factorization method, and the processed data is used as the score value of the sports user

(4) The adjacent optimizer consolidates the similarity between flow motions in two stages to optimize the neighborhood method

(5) Fuse the loss function and the results of the 4 steps according to the initial ratio

(6) Combine the results of the time loss coefficient with the dynamic merge module to generate a list of suggestions

(7) Dynamically train the weighting factor according to the call of the recommendation result to find the best factor value

Figure 1 shows the recommendation system architecture of the dynamic adaptive fusion algorithm.

Through calculation and analysis, the square root of these types of exercise sample sizes can reduce missing data and improve the accuracy of recommendations. The improved popular formula is as follows:

$$
\text{pop}_j = \begin{cases} \dfrac{|N(j)|}{\sum_{j \in pp} \|N(j)\|}, & \text{Popularity} \geqslant 200, \\[4mm] \dfrac{|N(j)|}{\sum_{j \in \text{pop} < 200} \sqrt{|N(j)|}}, & \text{Popularity} < 200. \end{cases} \tag{1}
$$
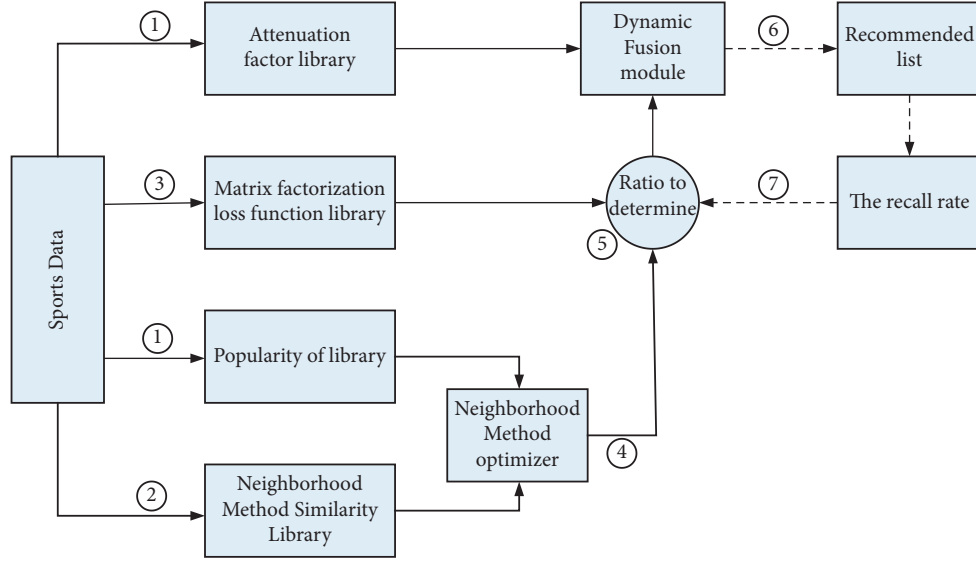
Figure 1: Dynamic adaptive fusion recommendation system framework.

All developments are time-sensitive, so this article introduces the concept of time fading. Suggestions for concave time decay function is as follows:

$$f(x) = \varphi^x. \tag{2}$$

Popular sports have long-term effects, while low-popular sports have temporary effects, so we first define the time frame function and use it as an exponential variable.

$$f_{ts} = -(T_c - T_0),$$
$$TD_j = e^{f_{ts}} \times \frac{1}{pop_j}. \tag{3}$$

The neighborhood method is used to calculate the similar relationship between users and sports according to formulas.

$$W_{ij} = \frac{|N(i) \cap N(j)|}{\sqrt{|N(i)\|N(j)|}}. \tag{4}$$

The matrix factorization technique uses the following formula to optimize the function of the power difference.

$$f_{loss} = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\mathbf{R} - \mathbf{u}_i^T \times \mathbf{v}_j \right)^2. \tag{5}$$

First, we define a similarity consistency function to ensure the consistency of the similarity matrix between users and sports obtained in the two calculations.

$$f_{smc-u} = \sum_{i=1}^{n} \left(\frac{\sum_{u_p \in KNN(u_i)} W_{ip} \times \mathbf{u}_p}{|KNN(\mathbf{u}_i)|} \right)^2,$$
$$f_{smc-v} = \sum_{j=1}^{m} \left(\mathbf{v}_j - \frac{\sum_{v_{jq}} \times \mathbf{v}_q}{|KNN(\mathbf{v}_j)|} \right)^2. \tag{6}$$

Whether users like sports and whether they are affected by popularity, so the optimization by multiplying the popular pop_j by the f_{smc-v} function is based on the proximity method.

$$f_{smc} = f_{smc-v} \times pop_j + f_{smc-u}. \tag{7}$$

Combining adaptive alpha weighting with time saving, used for time-based, adaptive f_{taf} mixing function, based on adjacent matrix transformation methods and techniques.

$$f_{taf} = (\gamma \times f_{smc} + (1 - \gamma) \times f_{loss}) \times TD_j. \tag{8}$$

Finally, we introduced a normalization function to prevent the dynamic adaptive recommendation algorithm from affecting the noisy data, so as to obtain the final objective function.

$$f_{MFDA} = f_{taf} + \lambda \times \left(\| u_i \|^2 + \| v_j \|^2 \right),$$
$$u_i = \frac{\text{Nei}(u) + (1 - \gamma) \times TD_j \times V \times R_i}{\gamma \times TD_j + (1 - \gamma) \times TD_j \times V \times V^T + \lambda},$$
$$v_j = \frac{\text{Nei}(v) \times pop_j + (1 - \gamma) \times TD_j \times U \times R_j}{\gamma \times TD_j \times pop_j + (1 - \gamma) \times TD_j \times U \times U^T + \lambda}. \tag{9}$$

Using the optimal solution of the alternating least square method, the prediction score matrix is finally obtained.

$$\text{Nei}(u) = \frac{\gamma \times TD_j \times W_{ip} \times u_p}{|KNN(u_i)|},$$
$$\text{Nei}(v) = \frac{\gamma \times TD_j \times W_{jq} \times v_q}{|KNN(v_j)|}. \tag{10}$$

In the simple addition fusion algorithm, the ratio of the neighbor method matrix decomposition method is 0.5. For comparison, the adaptive weighting factor of MFDA is fixed

TABLE 1: The relationship between times, weighting factors, and RMSE in different algorithms.

Index	10 to 29 years old		30 to 49 years old		50 to 69 years old	
	FSWA	MFDA	FSWA	MFDA	FSWA	MFDA
Training times weight factor RMSE	4	5	3	3	5	4
	0.4	0.15	0.2	0.34	0.7	0.74
	1.283	0.812	1.221	0.731	1.218	0.793

at 0.5. These three age records are applied to the FSWA and MFDA algorithms. The specific results are shown in Table 1.

3.2. Internet of Things Speech Recognition.

Sound waves are transmitted from the outer ear to the machine and converted into mechanical energy. When the last sogol (forged bone) in the cochlea moves, the flow of cochlear fluid generates traveling waves from the basement membrane. The impulse response of the cochlear basement membrane is represented by the function $\psi(t)$.

The response integral is zero.

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0. \tag{11}$$

Mean square integrable or finite energy.

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 dt < \infty. \tag{12}$$

Satisfy,

$$\int_{-\infty}^{+\infty} \frac{\Psi(\omega)|^2}{\omega} d\omega = C,$$

$$\Psi(\omega) = \int_{-\infty}^{+\infty} \psi(t)e^{-j\omega t}d\omega, \tag{13}$$

where $f(t)$ can all be mean square integral functions. The auditory conversion that converts the impulse response $\psi(t)$ of the base film to $f(t)$ is defined as follows:

$$T(a,b) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)dt. \tag{14}$$

The above formula can also be written as follows:

$$T(a,b) = \int_{-\infty}^{+\infty} f(t)\psi_{a,b}dt,$$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right). \tag{15}$$

This ensures that all values of α and b have the same energy.

$$\int_{-\infty}^{+\infty} \left|\psi_{a,b}(t)\right|^2 dt = \int_{-\infty}^{+\infty} |\psi(t)|^2 dt. \tag{16}$$

Definition formula of cochlear basilar membrane filter.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\left(\frac{t-b}{a}\right)^\alpha \exp\left(-2\pi f_L\beta\left(\frac{t-b}{a}\right)\right)$$

$$\cos\left(2\pi f_L\left(\frac{t-b}{a}\right) + \theta\right)u(t). \tag{17}$$

The demand for inverse conversion comes from the reconstruction of decomposed signals such as music synthesis or noise reduction. Similarly, the original signal can be restored if it is inversely transformed.

$$f(t) = \frac{1}{C}\int_{a=0}^{+\infty}\int_{b=0}^{+\infty}\frac{1}{|a|^2}T(a,b)\psi_{a,b}(t)dadb. \tag{18}$$

Discrete-time auditory positive transformation.

$$T[a_i,b] = \sum_{n=0}^{N} f[n]\frac{1}{\sqrt{|a_i|}}\psi\left[\frac{n-b}{a_i}\right]. \tag{19}$$

Discrete-time inverse auditory transformation.

$$\widetilde{f}[n] = \frac{1}{C}\sum_{a_i=a_1}^{o_k}\sum_{b=1}^{N}\frac{1}{|a_i|}T[a_i,b]\psi\left[\frac{n-b}{a_i}\right], \tag{20}$$

$$h(a,b) = [T(a,b)]^2 \forall a,b.$$

In the next step, the output $S(i,j)$ of each hair cell represents the relevant neural spike in the center frequency response of the current filter.

$$S(i,j) = \frac{1}{d_i}\sum_{b=1}^{l+d_-1} h(i,b), \quad l = 1, L, 2L, \cdots; \forall i, j. \tag{21}$$

The output of hair cells is converted into energy-perceived loudness through a nonlinear cube root loudness function.

$$y(i,j) = [S(i,j)]^{1/3}. \tag{22}$$

Table 2 shows the frequency division under the Bark scale:

Assuming that the original ignition is $x(n)$, $0 < n < N$, the auditory conversion coefficients $T(a,b)$ of m channels and N points after auditory conversion are obtained. $1 \le a \le m$, $1 \le b \le N$, where α represents different channels corresponding to different center frequencies and b represents different positions at different times, depending on the audio signal processing result.

4. Design and Practical Application of Mass Sports Data Monitoring System

4.1. Demand Analysis of Mass Sports Data Monitoring System.

Before starting the topic, we determined that the system can meet the needs of three users and offline surveys. (1) We can

TABLE 2: Bark scale filter center frequency division.

Filter number	Center frequency (Hz)
1	50
2	150
3	250
4	350
5	450
6	570
7	700
8	840
9	1000
10	1170
11	1370
12	1600
13	1850
14	2150
15	2500
16	2900
17	3400
18	4000

FIGURE 2: Client use case diagram.

monitor the user's exercise index in real time. (2) The user's exercise index can be judged. (3) We can provide healthy exercise guidance according to the changes of users' exercise parameters.

The client mainly refers to the Android application. After the user registers with the client, if the login is successful, the APP will enter the basic interface of the function. The main interface of the APP is divided into five functional modules in the setting module: exercise equipment connection, exercise indicator monitoring, self-health assessment, healthy exercise guide, and user's health file.

(1) Connection of exercise equipment: to monitor the flow indicator, the user must connect the customer to a smart portable exercise monitoring device. The user clicks the "Connect Device" button to add a device, and the client's Bluetooth must be turned on before the device can be bound.

(2) Monitoring of exercise indicators: the functional exercise monitoring module is a module for real-time monitoring of user exercise data. This module mainly displays the flow data of the portable exercise monitor on the smart terminal and records the heart rate, body temperature, pulse, exercise speed, weight, and other exercise indicators. Users can monitor their movements in real-time anytime and anywhere, which is useful for accurately understanding their movements. Client use case diagram as shown in Figure 2.

When the user clicks "monitor exercise," the interface will take us to the "exercise data view" page. We can see the movement trend chart from the zero point of the week to the present. The data is updated every minute and the latest data is displayed in digital form in the upper left corner of the trend graph. And when we select an indicator, the indicator name will turn blue, and the trend chart corresponding to the indicator will be displayed. When you move the
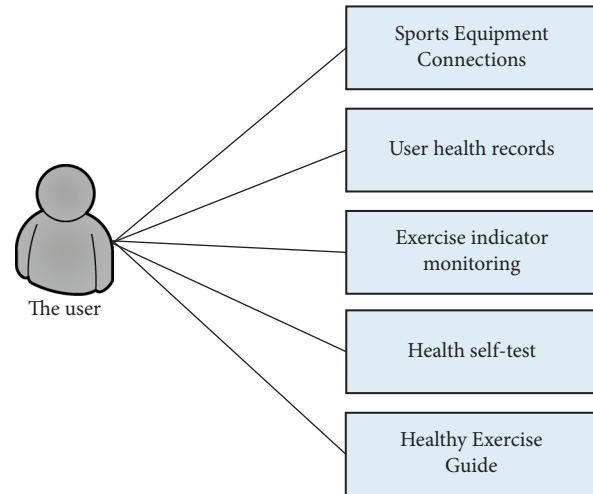
mouse to a specific point on the trend graph, the data floating layer of the selected indicator at that point will be displayed. The system will provide personalized health exercise guidance based on exercise data. As shown in Figure 3.

(3) Self-health assessment: the functional module consists of three parts: basic information, health status, and exercise habits. The basic information module records the user's gender, age, weight, and other information; exercise habits: the user's exercise intensity, daily exercise time, etc. If the information recorded in the user's health file is different from the user's actual information, you can manually change it.

(4) Healthy exercise guide: the system provides a way to obtain multiple sets of self-health questionnaires from various aspects. The system will also identify users' anxiety and depression levels based on user problems and provide them with detailed health improvement plans.

(5) User's health file: daily articles on various healthy exercise guidelines, healthy diet guidelines, etc., will be published on the APP homepage. Improve users' health awareness, develop a healthy lifestyle, and prevent chronic diseases at the source.

After the user completes the registration work from the mobile terminal, he can directly log in to the network terminal. If the user is not registered, the registration process must be completed, and the mobile terminal and the network terminal will synchronize user information. If the user is not logged in, he is not authorized to use some function modules. Users can input personal sports information through mobile devices or web devices to change their sports information. This information will be applied after the user clicks "Save," and the user's sports information on their mobile devices will be updated at the same time. If the practice data is different from the actual value, you can click the "Edit" or "Delete" button to perform the task.
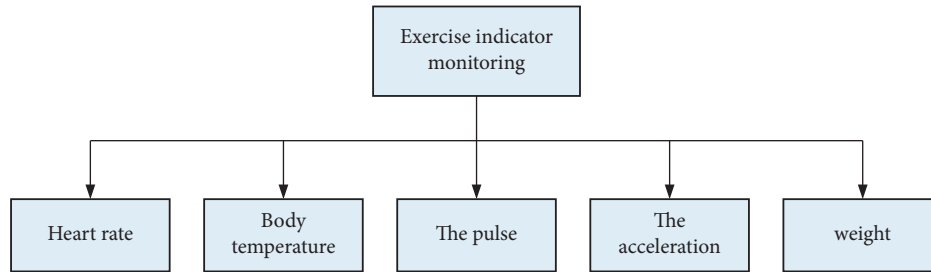
FIGURE 3: Use case diagram of sports indicator monitoring.

When we click on "Sports Trends," the trend chart of the day the user selects the indicator will be displayed by default. The trend chart can display up to two indicators at the same time. When we move the mouse to a specific point in time, we will see the data drawing layer, and the indicator is now selected. According to the user's health report, the system uses a simple algorithm to obtain the user's health score, evaluate the health level, provide health notifications, and personalize the user based on the level, and we provide a reasonable recovery set plan.

Customers of this system are assigned to Android smartphones with different user groups. The display that can carry the human motion parameters is easy to display and does not require complicated work. The web-side motion parameter management cloud platform is almost distributed on every home computer. Users can monitor their movement anytime and anywhere, and if they find an abnormality, they can stop the movement. The system uses MySQL (a scalable relational database to store data) and uses HTTP communication for data communication between the client and the server to ensure data accuracy. The wearable motion parameter monitor and the client can use the Socket data communication method to synchronize the motion index data in real time and download it to the client. Scalability refers to the data transmission capacity of the system, and the coordination of single or system-wide functional modules can meet the conversion needs. The system uses the highly scalable Android platform and JavaEE Web platform, and the database uses the relational database of MySQL and SQLite, so the system can perform read-write isolation and data fragmentation processing, which can improve the overall throughput.

4.2. System Architecture Design.

The system is based on the Android platform and configures wearable health monitoring through various sensors to collect the person's heart rate, temperature, pulse, acceleration, and other exercise parameters, and performs exercise detection through the Bluetooth module. The Android client allows users to view and evaluate their health level through the Web health management cloud platform to monitor their exercise status in real time. The system consists of three parts: wireless sensor module, intelligent terminal module, and server module. Among them, the wireless sensor module is a smart module and a Bluetooth module, the smart terminal module is a mobile phone and an Android cloud platform, and the server module is a background server of the database. The sports health monitor collects data about the user's exercise, and the server receives the traffic data measured by the sports health monitor, analyzes it, and responds to mobile phone customers' requests. The Android client and the practice display communicate with data via Bluetooth. After activating the display, it is waiting to establish a Bluetooth connection with the Android phone. When two Bluetooth devices are connected to the same Bluetooth device RFCOMM channel, the mobile phone and the monitor normally establish a Bluetooth connection, the mobile phone obtains the motion monitor data, the server receives the connection request from the client, and then establishes the interactive thread data logic with the client. The whole system architecture is shown in Figure 4.

After the portable human motion monitor communicates with the Android phone through the Bluetooth module and is properly configured through AT commands, the Android smart phone will collect traffic information on the motion monitor. The built-in Bluetooth portable exercise parameter monitor for mobile phones can communicate data through Socket. After the Android mobile phone receives the exercise instruction data, it will analyze the data, save it, and display it in a certain format. It is a waveform diagram of the Android mobile APP interface.

Android XML development use layout files to design the interface, and each functional module uses activity classes to interact with users. The system is mainly divided into a motion parameter module and a maintenance module. The motion control module includes a heart rate module, a body temperature module, a kinetic energy module, and a heart rate module. There are four service modules: save practice data, process data, completeness warning, and manage practice information, and then click practice information.

4.3. Functional Module Design

4.3.1. The Layout of the Registration Page.

When the user uses the system application for the first time, user registration is required. When designing the UI interface of the "user registration" functional module, first create the layout file register.xml corresponding to the Activity, and then define the relative layout on the other side. EditText controls and buttons will be created in this layout. The EditText control is used to enter user information. Click the "Button" button to send the registration information entered by the external user. Then,
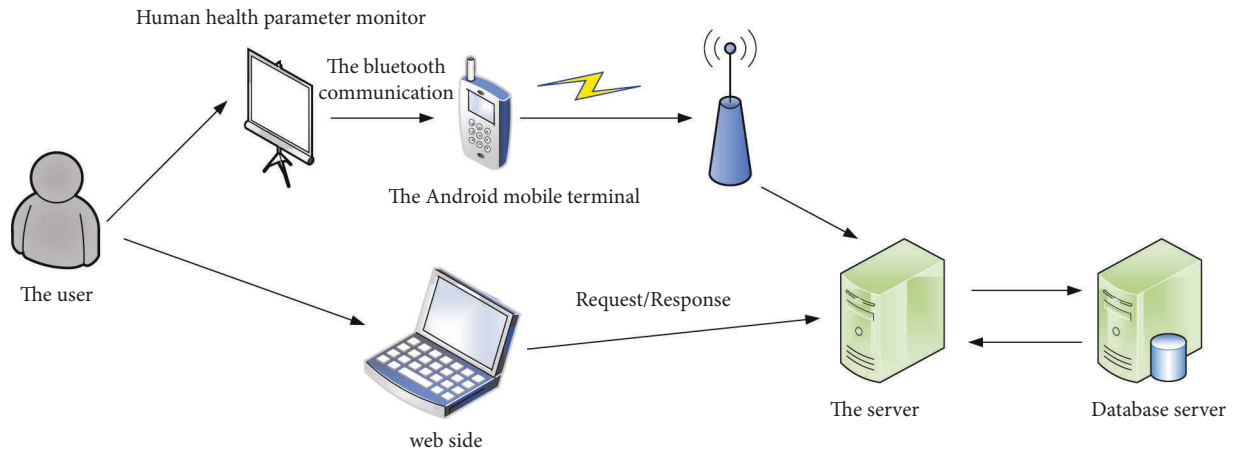
Figure 4: System overall architecture diagram.

generate another interface data.xml for receiving and sending data. The layout of the interface is relatively simple. Eight textViews have been added, and the registration form will display username, password, confirm password, gender, e-mail, and mobile phone. User information (such as b, Numbers, and real-time verification codes) will not be displayed in this interface of the application.

4.3.2. The Specific Realization of the Registration Function.

After creating the UI interface, create the RegisterActivity class inherited from the Activity class and implement the OnClickListener interface to cover onClick, where the getText method represents the registration information input by the user. Call the network interface class (requestapidat a) to create a getregistdata method, which interacts with the interface to create a user name, password, real name, gender, mobile phone number, e-mail, and other information entered by the user. Regardless of the login or registration function, it is very dangerous to send the password in plain text when connected to the network. The system uses the MD5 information digest algorithm combined with BASE64 to encrypt the password and calls the MD5 digest method to obtain a partial array of encrypted words.

This project is a network-side personal health exercise management that allows users to manage personal exercise parameters on various platforms (mobile and PC) to achieve personal health exercise management on multiple platforms. The system uses the open-source framework Spring, SpringMVC, and Mybatis to design a webchuk health exercise management cloud platform. The health exercise management cloud platform in the function module, registration module, and login module of the Weptuk system is divided into user health. The specific implementation of each functional module will be described in detail below.

If we are not familiar with the Webside health exercise management cloud platform, we need to register by clicking the register button at the top right of the homepage. When the user clicks the registration button, a registration mode dialog box will pop up, and the information that needs to be entered on the registration page includes user name,

password, gender, e-mail, and mobile phone number. The design of the registration interface mainly combines Spring mvc and JSP technology. In order to prevent unauthorized user input and malicious hacker attacks, the system introduces a verification mechanism, which regularly checks the information entered by the user on the register.jsp page. For example, when a user enters his mobile phone number on the registration page, if the input format or length does not match the regular expression format, the mobile phone number is entered incorrectly. At this time, a message will be displayed on the registration page showing the relevant mobile phone number and information that does not meet specifications. After entering the registered mobile phone number correctly, click "Get Confirmation Code," the page will display a 60-second countdown to start receiving the SMS verification code. The SMS verification function is an HTTP interface provided by the third party sendSms-ByHTTP. It is used to send Jixintong calls using the SMS platform Ali Da and sends the SMS template in the Ali-babaSmsUtils tool class, which sends short messages.

After the user successfully logs in, click "Healthy Sports File" to view past sports data. In this system, the exercise data of the health exercise management cloud platform is directly called from the database, and all these data are sent to the mobile client as exercise index data. The exercise index data is collected by the user using a wearable human exercise parameter monitor. The monitoring will be saved. The database uses paging technology to display a system generated from a large amount of relative exercise data measured at different time periods in a healthy exercise file.

4.4. Database Design.

The exercise data collected by the Android traffic monitor is retrieved via Bluetooth, and the exercise data is maintained locally and online to ensure that the data on the Android page is synchronized with the data on the server page. As shown in Table 3.

For Android, S'Lite is a lightweight database optimized for mobile devices. The system has developed an Android S'Lite database named joy Health, which contains five tables. (See Table 4)

TABLE 3: Mobile phone user information table (table name: user).

Field	Type of data	Field description
Id	Int	Data table auto-increment field
Username	Varchar	Registered user name
Password	Varchar	Password
Realname	Varchar	Actual name
Sex	Varchar	Gender
E-mail	Varchar	Register email
Phone	Varchar	Register mobile phone number
Date	Varchar	Registration time field
Is updated	Int	Whether to update to the server

TABLE 4: Heart rate measurement results of mobile phone users (table name: heart).

Field	Type of data	Field description
Heart_id	Int	Data table auto-increment field
Username	Varchar	Username field
Heart_value	Varchar	Weight measurement result field
Date	Varchar	Update time field
Is updated	Int	Whether to update to the server

5. Conclusion

In this course, we will design and implement a human health exercise management system based on Java Web and Android devices, and use cloud data processing technology to upload and save exercise parameters to the cloud. It is convenient for users to monitor exercise indicators in real time through the mobile client, and at the same time query exercise history data and compare exercise data through the network terminal, thereby improving exercise methods and exercise load, and avoiding physical discomfort caused by exercise. Scholars continue to study neural networks and audio signal processing. In this article, we use neural networks for speech recognition to conduct a more detailed study, so as to provide a preliminary exploration of the above problems. Deep learning is a new subfield of machine learning in the past few years, which focuses on NN model configuration and unsupervised learning using multiple layers of nodes. These detailed neural networks can better handle complex intelligence problems, and the way to process network model information is to imitate the human brain and can be used for speech recognition.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] W. Song, Z. Xiao, Y. Wang, and T. Chen, "Session-based social recommendation via dynamic graph attention networks," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pp. 555–563, Melbourne, Australia, February 2019.

[2] S. Ahmadian, M. Meghdadi, and M. Afsharchi, "A social recommendation method based on an adaptive neighbor selection mechanism," *Information Processing & Management*, vol. 54, no. 4, pp. 707–725, 2018.

[3] M. Benzeghiba, R. De Mori, O. Deroo et al., "Automatic speech recognition and speech variability: a review," *Speech Communication*, vol. 49, no. 10-11, pp. 763–786, 2007.

[4] W. Ping, K. Peng, K. Zhao, and Z. Song, "WaveFlow: a compact flow-based model for raw audio," in *Proceedings of the International Conference on Machine Learning*, pp. 7706–7716, Beijing China, September 2020.

[5] A. Madan and D. Gupta, "Speech feature extraction and classification: a comparative review," *International Journal of computer applications*, vol. 90, no. 9, pp. 20–25, 2014.

[6] B. Houlihan, "Sporting excellence, schools and sports development: the politics of crowded policy spaces," *European Physical Education Review*, vol. 6, no. 2, pp. 171–193, 2000.

[7] X. Yong, "Research on the evaluation of development level of principal sports industry based on analytic hierarchy process," in *Proceedings of the 2010 International Conference on E-Business and E-Government*, pp. 2949–2952, Guangzhou, China, July 2010.

[8] K. Yang, "The construction of sports culture industry growth forecast model based on big data," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 5–17, 2020.

[9] Y. N. Malek, A. Kharbouch, H. E. Khoukhi et al., "On the use of IoT and big data technologies for real-time monitoring and data processing," *Procedia Computer Science*, vol. 113, pp. 429–434, 2017.

[10] P. K. Dutta and S. Banerjee, "Monitoring of aerosol and other particulate matter in air using aerial monitored sensors and real time data monitoring and processing," *Journal of System and Management Sciences*, vol. 9, no. 2, pp. 104–113, 2019.

[11] Q. Zhang, "Evaluation and prediction of sports health literacy of college students based on artificial neural network," *Revista de Psicologia del Deporte*, vol. 30, no. 3, pp. 9–18, 2021.

[12] Y. T. Yue and Y. Yang, "Mobile intelligent terminal speaker identification for real-time monitoring system of sports training," *Evolutionary Intelligence*, vol. 23, pp. 1–12, 2020.

[13] Z. Wang and Z. Gao, "Analysis of real-time heartbeat monitoring using wearable device Internet of Things system in sports environment," *Computational Intelligence*, vol. 37, no. 3, pp. 1080–1097, 2021.

[14] J. Yang, J. Wen, B. Jiang, and H. Wang, "Blockchain-based sharing and tamper-proof framework of big data networking," *IEEE Network*, vol. 34, no. 4, pp. 62–67, 2020.

[15] M. Su, B. D. Fath, and Z. Yang, "Urban ecosystem health assessment: a review," *Science of the Total Environment*, vol. 408, no. 12, pp. 2425–2434, 2010.

[16] L. Savioja, V. Välimäki, and J. O. Smith, "Audio signal processing using graphics processing units," *Journal of the Audio Engineering Society*, vol. 59, no. 1/2, pp. 3–19, 2011.

Review Article

5G System Overview for Ongoing Smart Applications: Structure, Requirements, and Specifications

Hani Attar [iD],[1] Haitham Issa,[2] Jafar Ababneh [iD],[3] Mahdi Abbasi [iD],[4] Ahmed A. A. Solyman [iD],[5] Mohammad Khosravi [iD],[6] and Ramy Said Agieb[7]

[1]*Departement of Energy Engineering, Zarqa University, Zarqa, Jordan*
[2]*Departement of Electrical Engineering, Zarqa University, Zarqa, Jordan*
[3]*Abdul Aziz Al Ghurair School of Advanced Computing (ASAC), Luminus Technical University College, Amman, Jordan*
[4]*Department of Computer Engineering Engineering Faculty Bu-Ali Sina University, Hamedan, Iran*
[5]*Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Nişantaşı University, Istanbul 34398, Turkey*
[6]*Department of Computer Engineering, Persian Gulf University, Bushehr, Iran*
[7]*Department of Electrical Engineering, Faculty of Engineering, MTI University, Cairo, Egypt*

Correspondence should be addressed to Hani Attar; hattar@zu.edu.jo

Fifth-generation (5G) cellular networks are state-of-the-art wireless technologies revolutionizing all wireless systems. The fundamental goals of 5G are to increase network capacity, improve data rates, and reduce end-to-end latency. Therefore, 5G can support many devices connected to the Internet and realize the Internet of Things (IoT) vision. Though 5 G provides significant features for mobile wireless networks, some challenges still need to be addressed. Although 5 G offers valuable capabilities for mobile wireless networks, specific issues still need to be resolved. This article thoroughly introduces 5G technology, detailing its needs, infrastructure, features, and difficulties. In addition, it summarizes all the requirements and specifications of the 5G network based on the 3rd Generation Partnership Project (3GPP) Releases 15–17. Finally, this study discusses the key specifications challenges of 5G wireless networks.

1. Introduction

The tremendous expansion in mobile connections and innovative applications has raised the demand for new wireless services. This expansion elevates the demands for network technical indicators to be more rigorous. According to Ref. [1], about 20.4 billion Internet of Things (IoT) devices will be connected via machine to machine (M2M) by 2020. According to Ericsson [2], the number of linked devices is predicted to reach 75.4–100 billion by 2025. Meanwhile, according to Huawei, the fifth-generation (5G) subscribers will reach 2.8 billion by 2025 [3].

Furthermore, the emerging 5G technology is the foundation for many future technologies and applications. As a result, the 5G mobile communication system has strengthened at a vital juncture in history to usher in an era of comprehensive information and give a superior user experience. Recently, the industry has been conducting extensive research on crucial 5G technologies, particularly with the emergence of Industry 4.0. As a result, the related inspection, assessment, and verification system must be established and upgraded. With the increasing demand for mobile traffic and the emergence of several technologies that require very high data rates, very low latency, and a broad spectrum, the next generation, 5G, is required to address these issues.

5G technology has eight innovative features such as ultrafast up to 10 Gbps, ultralow latency (1 ms), high capacity (large bandwidth), numerous connected devices, constant availability, and coverage everywhere designed to provide low energy consumption and long battery life [4].

IoT technology is one of the significant drivers of 5G technology. 5G provides innovative infrastructure for the IoT, and IoT use cases that are expected to grow significantly in the future. As a result, the IoT will also make a big difference in communication technology design. While 5 G will bring dramatic changes to cellular systems and the IoT, it faces some significant challenges [5].

The improved massive MIMO technique is another key technology employed in 5G [6]. Employing numerous antennas at the transmitter and/or receive in massive MIMO will considerably improve the wireless system's spectral and energy efficiency [7].

Another approach that allows the antenna to direct radio waves in a particular direction and the receiver is beamforming. Beamforming boosted spectrum efficiency and reduced interference while simultaneously transferring more data from more antennas [8, 9]. Furthermore, recently developed waveforms such as FBMC, UFMC, F-OFDM, and GFDM can overcome pulse shaping, filtering, and precoding issues to reduce the out-of-band (OOB) leakage of OFDM signals [10]. All of these technologies, as well as a few others, are regarded as critical enablers of 5G networks.

However, some issues need to be addressed to implement 5G technology. For example, 5G needs to enable device-to-device communication (D2D) and intelligent vehicle services, provide a reliable network, and ensure privacy and security. While D2D communication improves cellular reuse and interference with femtocells and microcells, and D2D communication can affect system performance because of sharing the same radio resources in the D2D network [11]. In addition, 5G needs to ensure network reliability and availability anytime and anywhere. Security and privacy are two critical challenges that 5G must address. 5G security and privacy challenges increase as the number of users and data exchange increases, increasing the number of threats [12]. In addition, 5G needs to deliver error-free, high-precision data transmission to meet its needs and vision. This is a significant challenge, especially regarding user mobility, which produces a doubly dispersive fading channel that needs complex equalization [13]. All these issues must be addressed efficiently to meet the demand and reach the specific goals of 5G technology.

Several survey papers on 5G technology have been conducted and published in peer-reviewed journals. A comprehensive assessment of the evolution of cellular networks toward 5G networks has been published [14]. The authors emphasized the new architectural paradigm change in radio network layout, air interfaces, smart antennas, cloud, and heterogeneous radio access network design (RAN). They have also detailed the core physical layer technology in detail. Recent research on 5G and IoT has been reviewed in Ref. [15]. They analyzed the new requirements for 5G-enabled IoT. Next, they explained the main methods of 5G IoT and analyzed future IoT issues and trends.

Meanwhile, a study of 5G energy-efficient systems was conducted in Ref. [16]. This paper explored new paradigms, including New Radio (NR), nonorthogonal multiple access (NOMA), machine learning, and cashable networks.

Another research study reports energy-efficient scenarios for green communications, such as D2D communications, spectrum sharing, UltraDense networks (UDNs), Massive MIMO, mmWave networks, and IoT [17].

This paper conducts comprehensive research focused on 5G technology. First, the paper outlines 5G networks, including the history of cellular networks, 5G cellular network architectures, 5G service-based architectures, NR 5G technologies, and standalone (SA) and nonstandalone (NSA) modes. Next, this paper will focus on the main requirements and specifications of 5G technology. This includes high data rates, low latency, wide bandwidth, many connected devices, network availability, coverage, low energy consumption, and long life. Finally, 5G specifications challenges such as frequency band completion, large amounts of data, MIMO technology, beamforming, D2D communication, ultralow latency, ultrareliable networks, security, and privacy will be outlined. This paper will help researchers and developers interested in 5G technology understand the challenges of the leading 5G requirements and specifications.

The rest of the paper is structured as follows. Section 2 provides an overview of the 5G system. We reviewed the fundamental requirements of 5G in Section 3, which included the eight 5G enabling critical criteria. Following that, Section 4 outlines the most pressing 5G standard challenges. Finally, Section 5 concludes the paper.

2. 5G System Overview

The 5G mobile system is the most recent mobile communication network technology that provides a new way to communicate, including widespread connectivity. Compared to the preceding 1st generation (1G) to the 4th generation (4G), 5G offers much higher performance. Furthermore, 5G should enable a new type of connectivity and applications such as smart vehicles, transport and car communications, and massive video downloads, as media is needed everywhere, a considerable increase in human, IoT interaction, and remote control with haptic feedback, all requires widespread connectivity, as shown in Figure 1.

At the same time, it has been designed to deliver a very low data rate to meet the needs of various applications, such as sensors and IoT applications [18]. In other words, it supports a wide range of applications, from those requiring a low data rate to those requiring a very high data rate with minimal latency. Even though the 5G network initially relied on legacy networks, the 5G cellular architecture is regarded as a heterogeneous network that must comprise macrocells, microcells, picocells, and relays. 5G speed tops out at 10 gigabits per second (Gbps), which is 10 to 100× faster than what you can get with 4G, as shown in Figure 2.

Furthermore, the mobile small cell system is an essential component of the 5G wireless cellular system, which includes both mobile relay and small cell systems.

2.1. Architecture of the 5G Cellular Network. It is possible to notice from the existing 5G network that its multiple access methods remain stable, as evidenced by contemporary
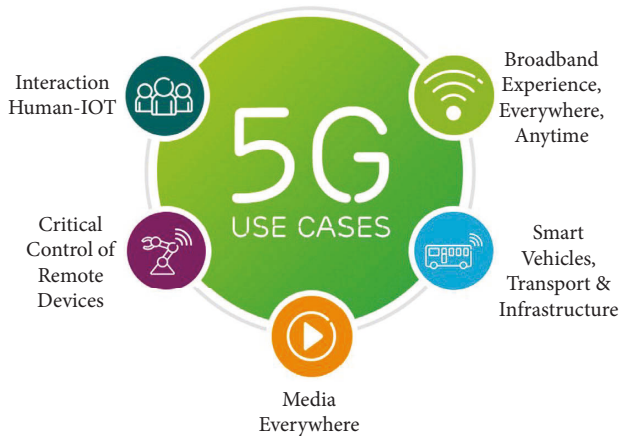
Figure 1: 5G use cases.
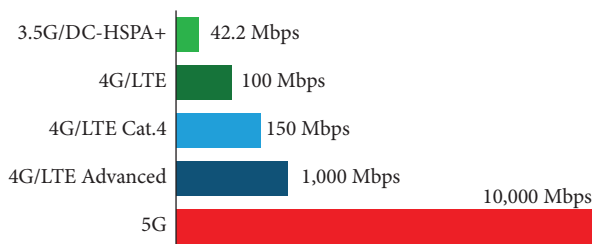
Figure 2: Comparison between 3G, 4G, and 5G data rates.

technologies such as orthogonal frequency-division multiple access (OFDMA), which will most likely be used for the next 50 years [19]. In addition, there is no need to adjust wireless settings from 1G to 4G. However, additional implementation improvements can be made to the base network to meet user criteria [20]. As a result, after the commercial development of 4G is completed, package providers will migrate easily to the 5G network. To address the issues of 5G systems and meet users' needs, developing a 5G wireless cellular architecture requires rapid development. In Ref. [21], most wireless users spent about 80% of their time indoors and 20% outdoors. According to contemporary wireless cellular architecture, an outdoor base station at the cell's center supports outdoor and interior user contact. As a result, the signals would be transferred around the inside walls, allowing interaction between indoor and outdoor users. As a result, significant penetration loss occurs, resulting in a decline in performance due to a fall in spectrum effectiveness and the energy efficacy of wireless communications. To address this challenge, a new design method for building the 5G cellular architecture for various outdoor and indoor scenarios has been developed [22]. Massive MIMO technology [23], which allows for the geographical distribution of tens or hundreds of antenna units, is one of the design strategies that would limit penetration through interior walls. Instead of the conventional MIMO systems that use two or four antennas, massive MIMO systems use big array antenna components that provide significant capacity benefits. For indoor coverage, multiple technologies such as millimeter-wave communication [24], small cells, Wi-Fi,

ultrawideband [25], and visible light communication [26] are suitable for short-range, high data rate responses. However, visible light and millimeter-wave responses require the utilization of higher frequencies, which are not typically used in cellular responses. Notably, long-distance and external deployment of these high-frequency waves is not advised because the waves would not successfully penetrate dense materials.

On the other hand, these waves might be efficiently disseminated by flora, gases, and raindrops. Millimeter-wave and visible light responses have a wide bandwidth and can increase the data transmission rate of indoor setups [27]. In addition to the emergence of new spectra that are generally not used for radio interactions, another approach to solving the problem of spectrum shortages is to use cognitive radio (CR) networks to improve the spectral utilization of new radio spectra [28]. The heterogeneity of 5G cellular architectures requires the integration of relays, small cells, microcells, and macrocells.

In particular, the mobile small cell idea is a crucial component of the 5G wireless cellular network, which comprises small cell concepts and mobile relays [29]. This concept was developed to preserve users' connectivity with significant mobility in high-speed trains and vehicles. Furthermore, mobile small cells are installed inside moving vehicles to connect with users, while the massive MIMO unit is installed on the vehicle's exterior to communicate with the outdoor base station. From the user's point of view, the mobile small cell is identified as a regular base station, and the associated user is identified as a single entity with a base station, indicating the separation of indoor and outdoor construction. Furthermore, small mobile cell users employ high data rates for bandwidth-intensive services while reducing signaling overhead. The 5 G wireless cellular network design has two logical layers: network cloud and radio network [30].

Different elements with different functions make up a wireless network. The network functions virtualization (NFV) cloud comprises a control plane entity (CPE) and a user plane entity (UPE), with the more critical utility of the planes associated with the control plane and user plane, respectively.

Notably, the 5G cellular network architecture is critical in both the backhaul and front-end networks. Figure 3 shows a multilayer network with macrocells covered by D2D links, femtocells, picocells, and relays. Implementing multiple layers in a cellular network architecture results in the excellent management of interlayer and intralayer interference, resulting in improved overall power consumption, spectral efficiency, coverage, and capacity.

2.2. New Radio 5G Technology. The 3GPP is in the process of defining a new 5G radio interface known as NR [31]. Enhanced Mobile Broadband (eMBB) will remain substantial, driving demand for increased system capacity, enhanced coverage, and faster data rates. Nonetheless, the ambitions of 5G transcend beyond eMBB. For example, MMTC, often known as IoT, is concerned with enabling meager device
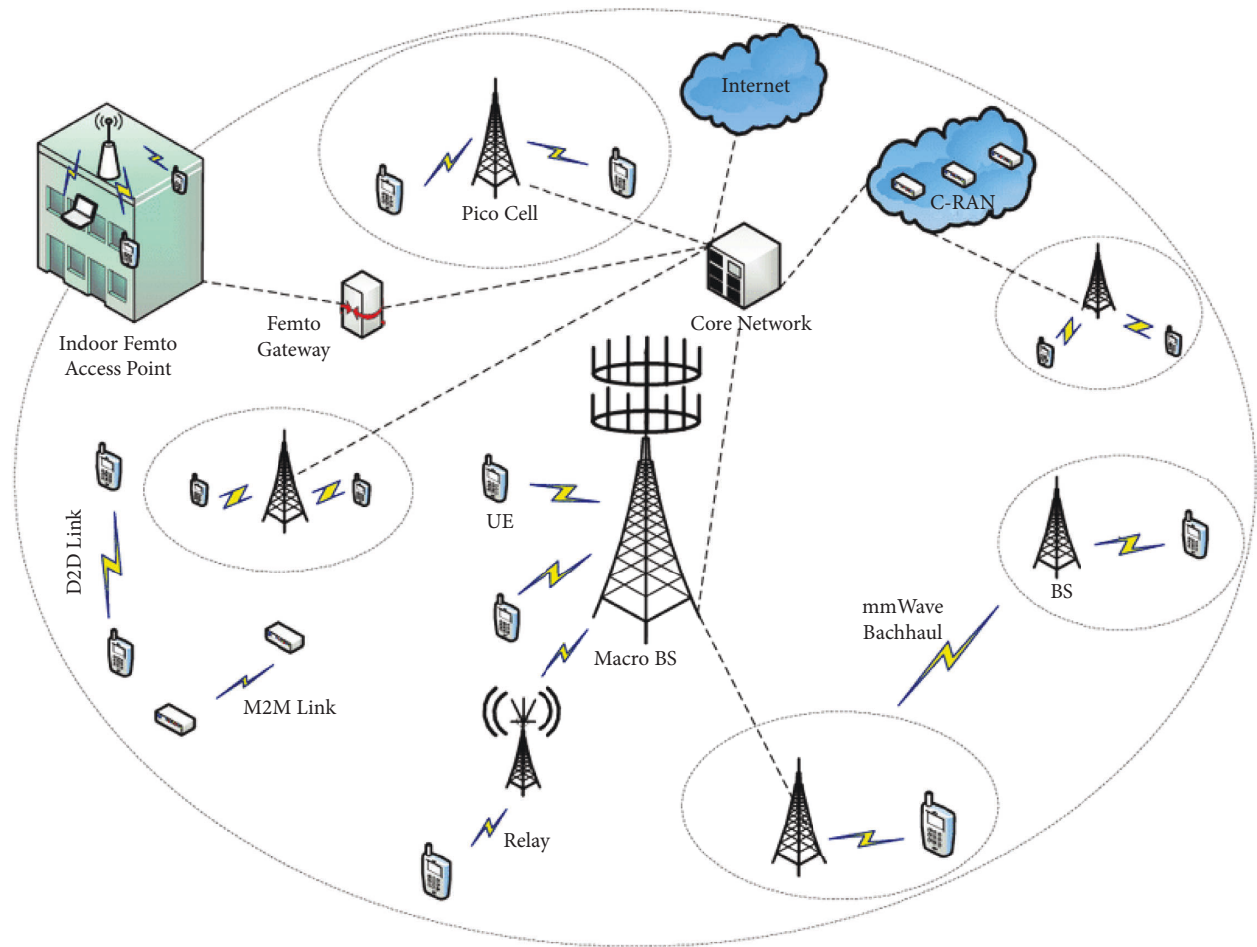
FIGURE 3: 5G cellular network architecture.

cost and energy consumption, offering extreme coverage, and handling a massive number of devices.

Another 5G goal is Ultra-Reliable Low-Latency Communication (URLLC), which provides data delivery with very low latency, for example, addressing critical industry applications. Furthermore, it is feasible that new and unexpected use cases will emerge during the life of NR. As a result, forward compatibility, which allows for the easy implementation of future optimizations, is a critical design requirement. Furthermore, improved network energy performance than the current systems is a key principle. Figure 4 depicts the several critical technologies that NR will support in 5G [32].

2.3. Standalone and Nonstandalone Mode. To improve 5G end-to-end performance, multiconnectivity is the optimal solution that combines long-term evolution (LTE) with reliable sub-6 GHz connections, such as simultaneous millimeter-wave connections. It can be used in the NSA mode with 5G NR [33]. This means that the 5G NR NSA mode means that the 4G network is used to connect to the control plane, and the 5G NR network is dedicated to the user plane. In contrast, in the standalone (SA) mode, both the control plane and the user plane use only the 5G NR leading network. Figure 5 shows both standalone and nonstandalone 5G NR systems. Technical work on NR began in the spring of 2016. The first release was part of Release 15 of 3GPP, and the NR specification was completed in late 2017. This initial release is limited to nonstandalone NR operations. NR devices rely on LTE for the initial access and mobility [34]. In addition, standalone NR operations are supported in the final release of 15 specifications available after June 2018 [35]. The difference between standalone and nonstandalone operations is primarily related to interface issues to the upper layers and core networks. The basic wireless technology is the same in both cases.

3. 5G Requirements

5G technology is driven by eight specification requirements, as shown in Figure 6:

(i) 100% coverage

(ii) Up to 10 Gbps data rate—>10 to 100x speed improvement over 4G and 4.5 G networks

(iii) one-millisecond latency

(iv) 1000 × bandwidth per unit area
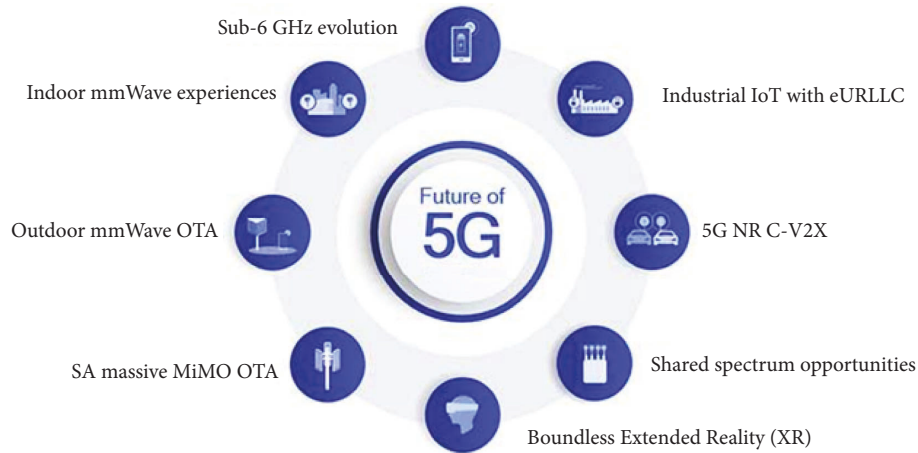
FIGURE 4: Technologies supported by 5G NR.
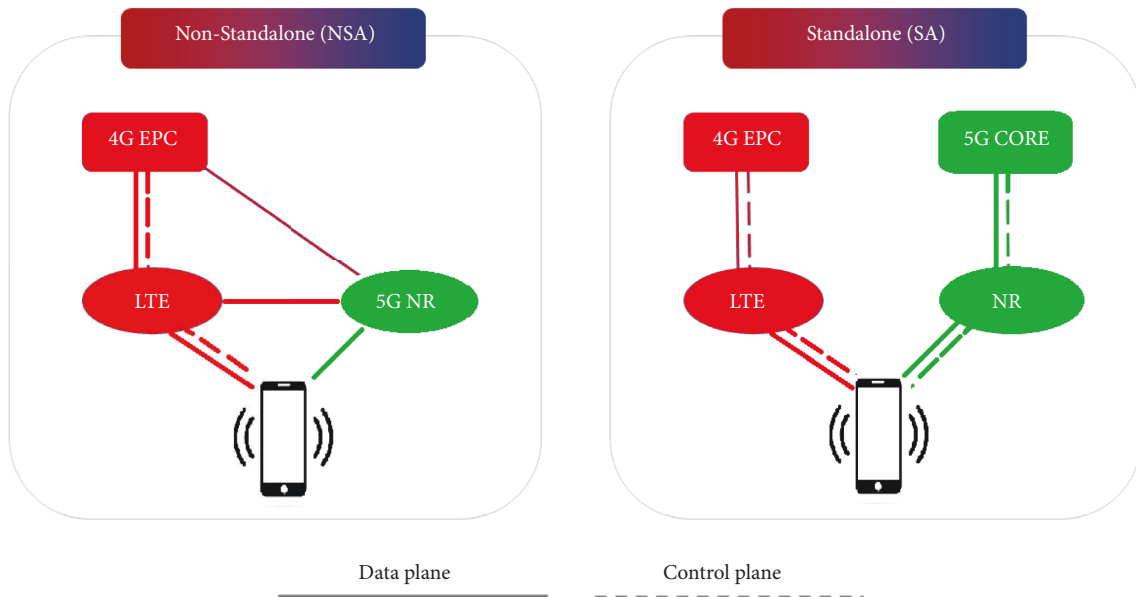
FIGURE 5: Nonstandalone 5G NR and standalone 5G NR.

(v) Up to 100x number of connected devices per unit area (compared with 4G LTE)

(vi) 99.999% availability

(vii) 90% reduction in network energy usage

(viii) Up to 10-year battery life for the low-power IoT device

3.1. High Coverage. Another issue for which 5G is looking for answers is coverage. 5G necessitates the perception of complete coverage [36]. By providing adequate coverage, technologies like IoT, D2D, and V2V can connect with the network from any location. Furthermore, because 5G enables network availability of up to 99.999 percent and coverage of nearly 100 percent, various services, including the technologies mentioned above, will be able to access the network at all times and from any location.

3.2. High Data Rate. The massive expansion in mobile connections over the last few years has increased the demand for large data volumes. This is the primary driving force behind the need for a high data rate in 5G technology. The 5G technology will support many services and applications that require high data rates. This system is designed to boost data rates up to 10 Gbps, a 10-fold improvement over the 4G network. Several advanced technical solutions, such as mmWave, massive MIMO, and various modulation and coding methods, enable 5G technology to achieve this goal [37].

3.3. Substantially Low Latency. One of the most significant needs that 5G focuses on is providing very low latency, where latency is the amount of time it takes for a request to complete an end-to-end round trip or the time it takes for the network to respond [38]. 5G latency will be significantly

1-millisecond latency

10 to 100x speed improvement over 4G

99.999% availability

1000x bandwidth per unit area

5G Technology

100% coverage

90% reduction in network energy usage

Up to 100x number of connected devices per unit area

10-year battery life for low power IoT device
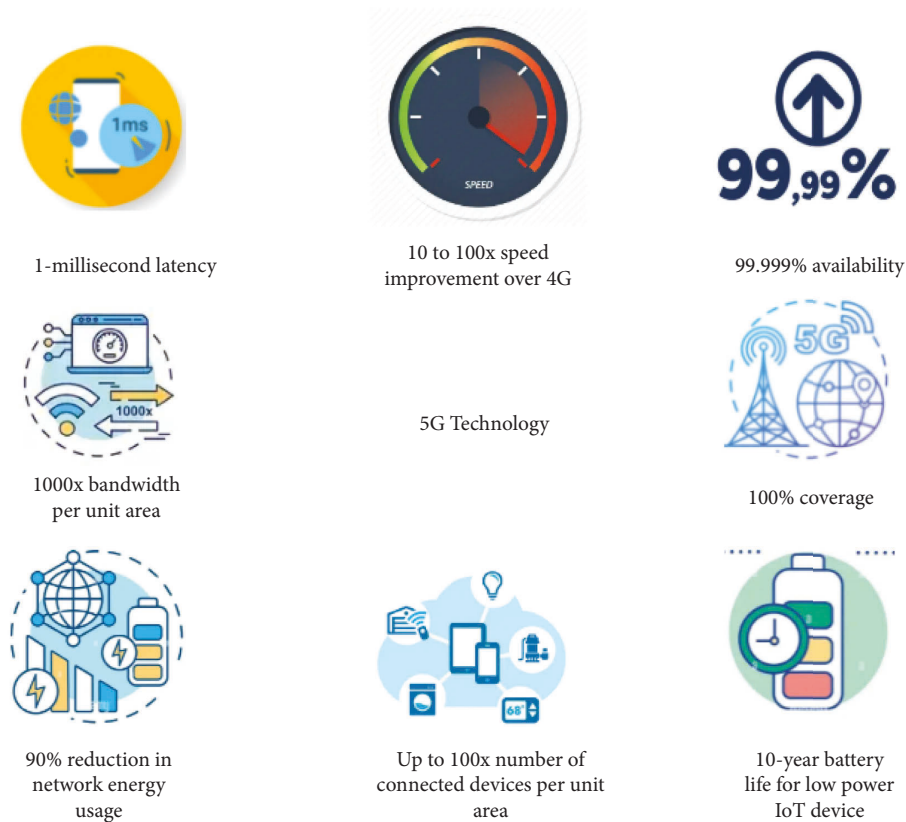
Figure 6: Specification requirements for 5G technology.

reduced to 1 ms end-to-end round trip. The end-to-end round trip delay latency of a 5G network is ten times lower than that of a 4G network [39]. The low latency of 5G will enable various applications, such as AR, VR, self-driving cars, and remote surgery. [40].

3.4. Wider Bandwidth. Increasing capacity causes an increase in bandwidth. The traditional LTE network operates on a frequency band of 3 GHz. This property restricts 5 Gbps implementations and complicates the design difficulty [41]. A high bandwidth per unit space enables a large number of linked devices in a small area. This is possible using mmWave technology [42]. 5G technology will give 1000 times the bandwidth per unit area while increasing capacity by up to 10 Gbps [43].

3.5. Massive Number of Connected Devices. Another significant aim of the 5G system is dramatically expanding capacity and enabling a massive number of concurrently connected devices. Compared to legacy systems, 5G will support up to 10–100 times the number of concurrently connected devices [44]. Massive IoT devices and sensors will be connected to the 5G network because it is a crucial infrastructure network for IoT. However, the communication channels must also ensure a very high data throughput and minimal latency. Although connecting billions of devices and sensors to the network simultaneously is difficult, this goal can be realized by utilizing solutions like massive MIMO, D2D technology, and cloud RAN [45].

3.6. High Availability. Many of the services and applications enabled by the 5G network necessitate constant network connectivity. As a result, 5G promises to deliver a perception of 99.999 percent availability, implying that the network will always be operational. This functionality would enable users and devices to connect to and access the network at any time [46].

3.7. Enhanced Battery Life. The battery life in 5G is expected to be 10 years. This goal is attainable with the advancement of battery technology and power-consuming electronics. Increasing battery life is critical for devices with limited battery capacities, such as laptops, cellphones, and tablets [47].

3.8. Low Energy Usage. 5G will lower network energy consumption by up to 90%. This reduction in energy consumption by 5G must occur as it provides extraordinarily high speed [31, 48]. Even though 5G provides fast data rates, low latency and coverage, and 99.999 percent availability, it will increase energy consumption by allowing rapid switching between sleep and active modes.

4. 5G Specifications Challenges

Because of the high demand and complex features of 5G networks, various technological challenges must be addressed.

4.1. Wide Frequency Bands. Network traffic and required data rates will increase in the future. These developments require a higher frequency band than that used in LTE systems. Therefore, the current frequency band with a range below 4 GHz cannot support 5G requirements. Therefore, we need new technologies that can support high traffic and data rates anywhere, under any conditions. Millimeter-wave technology is the best way to meet such needs at the same time. This is because mmWave provides wider frequency bands supporting 5G network traffic data.

Moreover, using the mmWave with massive MIMO and beamforming techniques presents suitable 5G network implementation options. However, supporting high traffic and additional frequency bands increase system complexity and thus make the 5G network application challenging. More issues of wide frequency bands are coverage limitation and higher attenuation compared to the low-frequency bands [49].

4.2. Huge Data Volume. Each year, the number of devices linked to the Internet grows considerably, as does the volume of data. The major challenge with 5G is the billions of gadgets and sensors connected to the Internet. According to Huawei and Information Handling Services (IHS) forecasts [50], the number of linked devices will reach 75–100 billion by 2025. As a result, 5G must be capable of handling vast amounts of data from an enormous number of connected devices and sensors and processing it concisely [51]. As data volume grows, it causes difficulty and complexity in securing, processing, regulating, and analyzing massive amounts of data [52].

4.3. Ultralow Latency. Providing an ultralow latency carrier is a severe situation confronted via way of means of 5G systems. Many crucial packages, including self-using and healthcare enterprise uses, require latencies of 1 ms. To offer an extensively low latency under 1 ms with huge statistics is a complicated challenge that 5G has to achieve. Besides that, machine-type communications (MTC) is a utility that 5G networks in all likelihood want to aid wherein the gadgets mechanically speak with every other possibility [53, 54]. For applications such as V2V communication, this type of communication requires very low latency. The METIS and METIS-II projects, which are European projects for traffic safety, and the "connected cars" use case addresses information exchange among vehicles and with the infrastructure to enable those as follows:

(i) A safer and more efficient transportation

(ii) Real-time remote computing for mobile terminals.

On the other hand, it has been proposed that traffic efficiency and safety be a typical application test case, with latency being critical in system evaluation. Intelligent traffic systems are a common scenario shown in METIS and METIS-II, in which cars require fast data sharing to avoid accidents. Low latency will also improve user experience in applications such as multiuser gaming. This reduction in latency necessitates technological innovation in waveform

design and flexible architecture in the network's higher layers, both of which can be met by wireless software-defined networking. [55, 56].

4.4. Ultrareliable Network. The accuracy of data transmission without errors is known as reliability. Several future applications and services will want highly dependable networks that will not tolerate failures. Operating at high frequencies using mmWave, 5G will make wireless technologies more vulnerable. As a result, 5G must address this difficulty by utilizing technologies such as beamforming, MIMO, and software-defined networking (SDN) [57].

4.5. Security and Privacy. In wireless communication systems, security and privacy are two of the most critical considerations. They are also the two most significant problems for the 5G system. 5G will accommodate many connected devices with fast data rates and massive capacity, raising security and privacy concerns. As a result, 5G must provide and guarantee end-to-end privacy and security for users. The vast number of connected devices to the 5G network with massive amounts of data and security and privacy concerns were two of the critical issues that 5G faced. To protect data and users' privacy, security and privacy must be considered [58, 59].

4.6. Smart Automobiles. Once 5G is built, self-driving automobiles and smart automobile technologies will be available. Smart cities are also projected to be launched to support this application. 5G will enable smart vehicles to connect and serve as a data transfer hub for communication with roadside infrastructure and other services [60]. Because of safety-critical use cases and rapidly changing vehicular network architecture, enabling smart automotive technologies necessitates networks with significantly reduced latency and exceptionally high dependability [13]. As a result, the 5G network must be designed to meet the rigorous QoS requirements of smart vehicle systems [54].

4.7. Spectrum Availability. The frequency spectrum is rare and is currently a supersaturated product. Therefore, it is important to evaluate the availability of the spectrum. Existing mobile technologies use frequencies ranging from 300 MHz to 3 GHz [61]. Regardless of this, new 5G frequency bands are being investigated. 5G uses low and high frequencies, as well as very high frequencies. Low frequencies (<1 GHz) and high frequencies (1–6 GHz) are commonly used for mobile communications. On the other hand, very high frequencies (30–300 GHz) represent a new spectral option for 5G. In Europe, Japan, South Korea, and the United States, intervals between 3.5 and 4.5, and also 28 GHz include targeted or allocated licensed and unlicensed bands. In addition, the 39 GHz band is typically used for 5G in the United States and Canada. The 39 GHz band can be helpful as this bandwidth is unsaturated and therefore is available for mobile communications. Figure 7 shows 5G radio frequency spectrum [62].
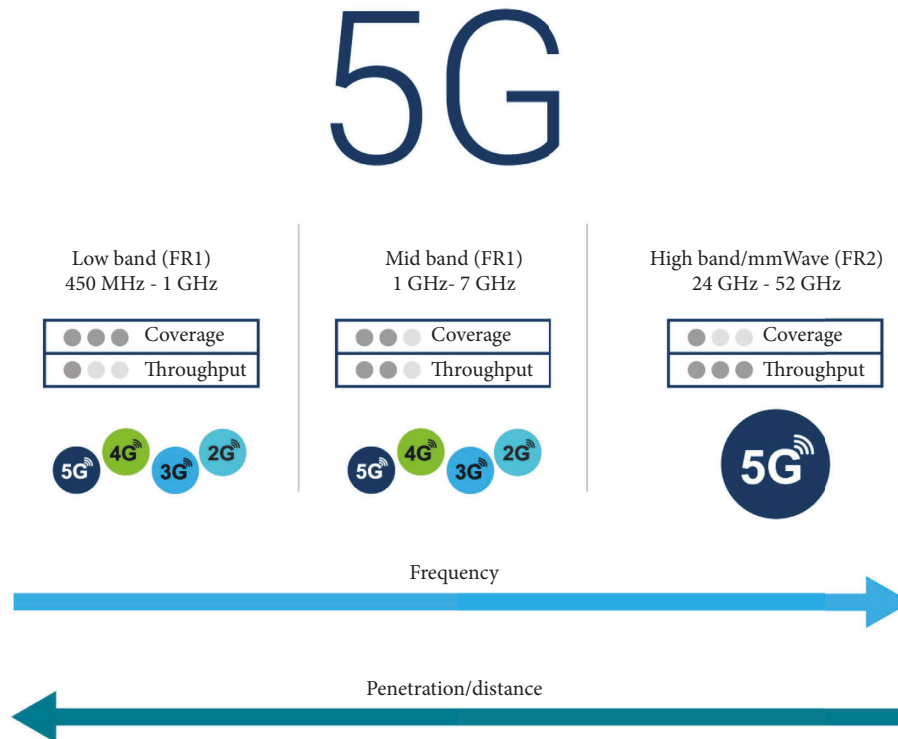
FIGURE 7: 5G radio frequency spectrum.

4.8. Air Interface. Due to the exacting requirements of 5G, a new air interface must be developed. Thus, new air interfaces, known as NR, are being created according to modulation schemes, which are being developed based on the scope of future mobile communication systems [63] since 5G using mmWaves technology, having small wavelengths, will require small antenna sizes. In other words, 5G uses a massive of small antennas. Exploiting beamforming technology directs the antenna's signals into a specific and desired direction, enhances the data rate, and increases the signal strengths with low interference [64]. However, 5G must overcome hardware design challenges, power consumption, directional accuracy, and analog-digital and digital-analog processes [65].

4.9. Special Protocols Adapt with 5G. Recent research is directed to apply particular protocols to improve the 5G wireless network specifications, such as the data rate, network congestion, and security. Applying network coding, co-operation networks, hybrid networks, etc. is an important technique to achieve the desired improvement.

Accordingly, applying network coding and co-operation networks resulted in what is called co-operation network coding (CoNC), which resulted in improving the data rate, network congestion, security, transmission packet error rate (PER), and bit error rate (BER) [66–69], where [51] the hybrid network was applied to improve the transmission range and to improve the PER. Moreover, improving the application over the physical layer was widely proposed

[70–72], and 5G applications more and more cover medical fields.

Future work is planned so as to focus on viewing an inclusive article that accumulates information for the medical applications that are suggested for 5G and B5G based on our published work in [73–77].

5. Conclusion

A complete survey of 5G technologies requirements and specifications challenges has been undertaken in this study. The survey describes the overall network architecture and essential requirements for the 5G network. Meanwhile, the paper introduced technologies that would make 5G a reality, including mmWave, massive MIMO, beamforming techniques, and other advanced technologies. Finally, the study outlined several essential difficulties that must be effectively solved before 5G technology can be globalized. This survey paper provides readers with a concise and in-depth review of 5G wireless networks.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5G network with the advent of 6G: architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67512–67547, 2021.

[2] Ericsson, *Leveraging the Potential of 5G Millimeter Wave*China, 2021, https://www.ericsson.com/en/reports-and-papers/further-insights/leveraging-the-potential-of-5g-millimeter-wave.

[3] Ericsson, "Ericsson mobility report," 2018, https://www.ericsson.com/491e17/assets/local/reports-papers/mobility-report/documents/2018/ericsson-mobility-report-june-2018.pdf.

[4] A. Morgado, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "A survey of 5G technologies: regulatory, standardization and industrial perspectives," *Digital Communications and Networks*, vol. 4, no. 2, pp. 87–97, 2018.

[5] A. Gupta and M. Gupta, "A review on activities of fifth generation mobile communication system," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 1125–1135, 2018.

[6] R. Akl and R. Akl, "Massive MIMO systems for 5G and beyond networks-overview, recent trends, challenges, and future research direction," *Sensors*, vol. 20, no. 10, pp. 2753–2835, 2020.

[7] M. Al-Rawi and M. Al-Rawi, "Spectral and energy efficiency of massive MIMO system," *International Review of Applied Sciences and Engineering*, vol. 7, no. 2, pp. 71–75, 2016.

[8] A. Saraereh and A. Ali, "Beamforming performance analysis of millimeter-wave 5G wireless networks," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 5383–5397, 2022.

[9] M. Moubadir, A. Mchbal, N. A. Touhami, and M. Aghoutane, "A switched beamforming network for 5G modern wireless communications applications," *Procedia Manufacturing*, vol. 32, pp. 753–761, 2019.

[10] A. J. Tiwana, M. Zeeshan, T. Ashraf, M. U. Farooq, K. Shahzad, and A. Akhunzada, "Dynamic link adaptation for filterband multicarrier in networks with diverse service quality and throughput requirements," *Telecommunication Systems*, vol. 79, no. 1, pp. 109–122, 2022.

[11] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, S. Li, and G. Feng, "Device-to-device communications in cellular networks," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 49–55, 2014.

[12] R. Khan, P. Kumar, D. N. K. Jayakody, and M. Liyanage, "A survey on security and privacy of 5G technologies: potential solutions, recent advancements, and future directions," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 196–248, 2020.

[13] H. H. Solyman, M. Khosravi, M. Bashir, A. K. Tavallali, and P. Tavallali, "Efficient equalisers for OFDM and DFrFT-OCDM multicarrier systems in mobile E-health video broadcasting with machine learning perspectives," *Physical Communication*, vol. 42, Article ID 101173, 2020.

[14] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[15] S. Li, L. D. Xu, and S. Zhao, "5G Internet of Things: a survey," *Journal of Industrial Information Integration*, vol. 10, pp. 1–9, 2018.

[16] M. A. Kumaraswamy and H. V. Kumaraswamy, "Energy efficient 5G networks: techniques and challenges," in *Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1317–1322, Sep, NY City, June 2020.

[17] P. Gandotra, R. J.-J. Suttizen, and C. Applications, *A Survey on green Communication and Security Challenges in 5G Wireless Communication Networks*Elsevier, Amsterdam, Netherlands, 2017, https://www.sciencedirect.com/science/article/pii/S1084804517302266?casa_token=7UYwFvm0_AIAAAAA.

[18] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, "5G evolution: a view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, pp. 127639–127651, 2019.

[19] P. Shah, S. Qureshi, and R. Butt, *Design and Analysis of 5G Network Architecture with Orthogonal Frequency Division Multiple Access Based Passive Optical Network*Elsevier, Amsterdam, Netherlands, 2021, https://www.sciencedirect.com/science/article/pii/S1068520021002273?casa_token=XJpne8HRvasAAAAA:84sntFuGUzXlAK38bWg7QOb6Wp3bUodR_gpNf9RVAS3sM4QLfvW5-QzsZXDx83blZL--Kt4qZg.

[20] A. Gupta, "A survey of 5G network: architecture and emerging technologies," *Recent Advances in Software Defined Networking for 5G Networks*, vol. 3, 2015.

[21] M. A. M. Albreem, "5G wireless communication systems: vision and challenges," in *Proceedings of the 2015 International Conference on Computer, Communications, and Control Technology (I4CT)*, p. 493, Beijing China, August. 2015.

[22] M. Tataria, R. Sun, W. Tufvesson, S. Kitao, and K. Kitao, "Microwave vs. Millimeter-wave propagation channels: key differences and impact on 5G cellular systems," *IEEE Communications Magazine*, vol. 56, no. 12, pp. 14–20, Dec. 2018.

[23] K. Nakagawa, I. Iura, N. Okazaki, M. Suyama, A. Okumura, and Y. Okumura, "Indoor experimental trial on hybrid 16-beam spatial-multiplexing for high SHF wide-band massive MIMO in 5G," in *Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pp. 7–11, NY City, August 2018.

[24] M. Ghaddar, I. Ben Mabrouk, M. Nedil, K. Hettak, and L. Talbi, "Deterministic modeling of 5g millimeter-wave communication in an underground mine tunnel," *IEEE Access*, vol. 7, pp. 116519–116528, 2019.

[25] X. Zhang, Y. Li, and W. Wang, "Ultra-wideband 8-port MIMO antenna array for 5G metal-frame smartphones," *IEEE Access*, vol. 7, 2022.

[26] S. Idris and U. Mohammed, "Visible light communication: a potential 5G and beyond communication technology," in *Proceedings of the 2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*, Abuja, Nigeria, December 2019.

[27] R. Faruque and S. Faruque, "An enhanced indoor visible light communication physical-layer security scheme for 5G networks: survey, security challenges, and channel analysis secrecy performance," *International Journal of Communication Systems*, vol. 34, no. 4, 2021.

[28] F. Hu and B. Chen, "Full spectrum sharing in cognitive radio networks toward 5G: a survey," *IEEE Access*, vol. 6, 2018.

[29] F. Tseng, H. Chao, and J. W. I. W. Communications, "Ultra-dense small cell planning using cognitive radio network toward 5G," *IEEE Wireless Communications*, vol. 22, 2015.

[30] J. G.-P. Chen and M. Xiao, "Computing and undefined," in *Interrelationship between Energy Efficiency and Spectral Efficiency in Cognitive Femtocell Networks: A Survey*Elsevier, Amsterdam, Netherlands, 2019.

[31] Y. Li, M. Chen, J. Xu, and L. Tian, "Power saving techniques for 5G and beyond," *IEEE Access*, vol. 8, 2020.

[32] Qualcomm Technologies, "Driving 5G NR technology evolution forward to unlock the full potential of 5G," 2022, https://www.qualcomm.com/news/onq/2019/02/19/driving-5g-nr-technology-evolution-forward-unlock-full-potential-5g.

[33] H. Bagheri and M. Noor-A-Rahim, "5G NR-V2X: toward connected and cooperative autonomous driving," *IEEE Communications Standards Magazine*, vol. 5, 2021.

[34] H. Attar, L. Stankovic, M. Alhihi, and A. Ameen, "Deterministic network coding over long term evaluation advance communication system," in *Proceedings of the 2014 Fourth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, pp. 56–61, Beijing China, June 2014.

[35] A. Hoglund, D. P. Van, T. Tirronen, O. Liberg, Y. Sui, and E. A. Yavuz, "3GPP release 15 early data transmission," *IEEE Communications Standards Magazine*, vol. 2, no. 2, pp. 90–96, 2018.

[36] M. Ahamed and S. Faruque, "5G network coverage planning and analysis of the deployment challenges," *Sensors*, vol. 21, no. 19, p. 6608, 2021.

[37] S. Baig, H. Asif, T. Umer, and S. Mumtaz, "High data rate discrete wavelet transform-based PLC-VLC design for 5G communication systems," *IEEE Access*, vol. 6, 2018.

[38] I. Parvez and A. Rahmati, "A survey on low latency towards 5G: RAN, core network and caching solutions," *IEEE Communications Surveys & Tutorials*, vol. 22, 2018.

[39] C. Li, J. Jiang, and W. Chen, "5G ultra-reliable and low-latency systems design," in *Proceedings of the 2017 European Conference on Networks and Communications (EuCNC)*, Oulu, Finland, July 2017.

[40] M. Lema, A. Laya, and T. Mahmoodi, "Business case and technology analysis for 5G low latency applications," *IEEE Access*, vol. 5, 2017.

[41] C. Ding, H. Sun, and H. Zhu, "Achieving wider bandwidth with full-wavelength dipoles for 5G base stations," *IEEE Transactions on Antennas and Propagation*, vol. 68, 2019.

[42] I. A. Arafah and M. A. Arafah, "Enhancing 5G small cell selection: a neural network and IoV-based approach," *Sensors*, vol. 21, no. 19, p. 6361, 2021.

[43] G. Ancans, A. Stafecka, V. Bobrovs, A. Ancans, and J. Caiko, "Analysis of characteristics and requirements for 5G mobile communication systems," *Latvian Journal of Physics and Technical Sciences*, vol. 54, no. 4, pp. 69–78, 2017.

[44] N. Al-Falahy, "Technologies for 5G networks: challenges and opportunities," *IT Professional*, vol. 19, 2017.

[45] A. K. Bashir, B. Arul, J. Raja, and N. M. F. Qureshi, "An optimal multitier resource allocation of cloud RAN in 5G using machine learning," *Transactions on Emerging Telecommunications Technologies*, vol. 30, no. 8, 2019.

[46] D. Oehmann and A. Awada, "Modeling and analysis of intra-frequency multi-connectivity for high availability in 5G," in *Proceedings of the 2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Porto, Portugal, July 2018.

[47] T. Dilemma, "More base stations, more antennas—less energy? - IEEE spectrum," 2022, https://spectrum.ieee.org/will-increased-energy-consumption-be-the-achilles-heel-of-5g-networks.

[48] M. Lauridsen, "Sleep modes for enhanced battery life of 5G mobile terminals," in *Proceedings of the 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, China, May 2016.

[49] J. Lee, E. Tejedor, and K. Ranta-aho, "Spectrum for 5G: global status, challenges, and enabling technologies," *IEEE Communications Magazine*, vol. 55, 2018.

[50] M. Zenglein and A. H.-M. papers, *Evolving Made in China 2025*kritisches-netzwerk.de, China, on China, and undefined 2019, 2019.

[51] H. H. Attar, A. A. A. Solyman, A. Alrosan, C. Chakraborty, and M. R. Khosravi, "Deterministic cooperative hybrid ring-mesh network coding for big data transmission over lossy channels in 5G networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 14, pp. 1–18, 2021.

[52] K. Z.-M. Microsystem, "Microsystems and Undefined," in *Sports and Health Big Data System Based on 5G Network and Internet of Things System*Elsevier, Amsterdam, Netherlands, 2021.

[53] E. Dutkiewicz, X. Costa-Perez, and I. K.-I. Network, *Massive machine-type communications*Seoul China, 2017.

[54] Z. Dawy, W. Saad, and A. Ghosh, "Communications, and Undefined," *Toward massive machine type cellular communications*, vol. 256, 2019.

[55] M. Siddiqi, H. Yu, J. J. Joung, and undefined, "5G ultra-reliable low-latency communication implementation challenges and operational issues with IoT devices," *Electronics*, 2019.

[56] M. A. Siddiqi, H. Yu, and J. Joung, "5G ultra-reliable low-latency communication implementation challenges and operational issues with IoT devices," *Electronics*, vol. 8, no. 9, p. 981, 2019.

[57] N. Zhang, P. Yang, S. Zhang, and D. Chen, "Software defined networking enabled wireless network virtualization: challenges and solutions," *IEEE Network*, vol. 31, 2017.

[58] R. Khan and P. Kumar, "A survey on security and privacy of 5G technologies: potential solutions, recent advancements, and future directions," *IEEE Communications Surveys & Tutorials*, vol. 22, 2019.

[59] X. Ji and X. Xiao, *Overview of 5G Security Technology*Springer, Heidelberg, Germany, 2022.

[60] U. Rajkumar and K. Rajkumar, "Background and research challenges for blockchain-driven 5G IoT-enabled industrial automation," *Blockchain for 5G-Enabled IoT*, vol. 23, pp. 33–59, 2021.

[61] M. Matinmikko-Blue, S. Yrjölä, V. Seppänen, P. Ahokangas, H. Hämmäinen, and M. Latva-Aho, "Analysis of spectrum valuation approaches: the viewpoint of local 5G networks in shared spectrum bands ieeexplore.ieee.org," in *Proceedings of the 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Seoul, Korea (South), October 2022.

[62] EXFO, "5G RF and 5G NR | Top 5 questions answered | EXFO," 2022, https://www.exfo.com/es/recursos/blog/rf-5g-new-radio-top-5-questions/.

[63] I. Da Silva, G. Mildh, and J. Rune, "Tight integration of new 5G air interface and LTE to fulfill 5G requirements," in *Proceedings of the 2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow, July 2022.

[64] F. Schaich and J. Chen, "FANTASTIC-5G: 5G-PPP project on 5G air interface below 6 GHz," 2022, http://fantastic5g.com/wp-content/uploads/2015/07/EuCNC-FANTASTIC-5G_final.pdf.

[65] G. Murugesan and G. Murugesan, "A comprehensive survey on air-interfaces for 5G and beyond," in *Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kanpur, India, June 2019.

[66] H. H. Attar, A. A. A. Solyman, M. R. Khosravi, L. Qi, M. Alhihi, and P. Tavallali, "Bit and Packet Error Rate evaluations for Half-Cycle stage cooperation on 6G wireless networks," *Physical Communication*, vol. 44, Article ID 101249, 2021.

[67] H. Attar, M. Alhihi, B. Stankovic, and L. Stankovic, "Network coding hard and soft decision behavior over the physical payer

using PUMTC," in *Proceedings of the 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pp. 471–474, Paris, France, June 2018.

[68] H. Attar, L. Stankovic, M. Ameen, and A. Ameen, "Deterministic network coding over long term evaluation advance communication system," in *Proceedings of the 2014 Fourth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, pp. 56–61, Korea China, June 2014.

[69] H. Attar, L. Stankovic, and V. Stankovic, "Cooperative network-coding system for wireless sensor networks," *IET Communications*, vol. 6, no. 3, pp. 344–352, 2012.

[70] H. Attar, "Efficient physical layer techniques for healthcare applications: Co-operative network coding algorithms and modified equalizers," *Intelligent Healthcare*, vol. 18, pp. 423–454, 2022.

[71] A. L. Michala, H. Attar, and I. Vourganas, "Secure data transfer and provenance for distributed healthcare," *Intelligent Healthcare*, vol. 241, p. 260, 2022.

[72] I. Vourganas, H. Attar, and A. L. Michala, "Accountable, responsible, transparent artificial intelligence in ambient intelligence systems for healthcare," *Intelligent Healthcare*, vol. 87, p. 111, 2022.

[73] E. Hammam, H. A. Attar, A. Amer et al., "Prediction of wear rates of UHMWPE bearing in hip joint prosthesis with support vector model and grey wolf optimization," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 6548800, pp. 1–16, 2022.

[74] A. Deif, H. A. Attar, A. Amer et al., "A new feature selection method based on hybrid approach for colorectal cancer histology classification," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7614264, pp. 1–14, 2022.

[75] K. Rezaee, G. Jeon, M. R. Khosravi, H. H. Attar, and A. Sabzevari, "Deep learning-based microarray cancer classification and ensemble gene selection approach," *IET Systems Biology*, vol. 16, no. 3-4, pp. 120–131, 2022.

[76] H. Attar, M. R. Khosravi, G. Igorovich, and M. Alhihi, "E-health communication system with multiservice data traffic evaluation based on a G/G/1 analysis method," *Current Signal Transduction Therapy*, vol. 16, no. 2, pp. 115–121, 2021.

[77] H. H. Attar, A. A. A. Solyman, A. E. F. Mohamed et al., "Efficient equalisers for OFDM and DFrFT-OCDM multi-carrier systems in mobile E-health video broadcasting with machine learning perspectives," *Physical Communication*, vol. 42, Article ID 101173, 2020.

Research Article

Virtual Reality and Internet of Things-Based Music Online Learning via the Graph Neural Network

Jian Lian [ID],[1] **Yanan Zhou** [ID],[2] **Lina Han** [ID],[3] **and Zhiguo Yu** [ID][1]

[1]*School of Intelligence Engineering, Shandong Management University, Jinan 250357, China*
[2]*Beijing College of Foreign Studies University, Beijing 100089, China*
[3]*Department of Electrical and Automation, Shandong Labor Vocational and Technical College, Jinan 250022, China*

Correspondence should be addressed to Yanan Zhou; sopronozhou@163.com

Virtual reality and the Internet of Things have shown their capability in a variety of tasks. However, their availability in online learning remains an unresolved issue. To bridge this gap, we propose a virtual reality and Internet of Things-based pipeline for online music learning. The one graph network is used to generate an automated evaluation of learning performance which traditionally was given by the teachers. To be specific, a graph neural network-based algorithm is employed to identify the real-time status of each student within an online class. In the proposed algorithm, the characteristics of each student collected from the multisensors deployed on their bodies are taken as the input feature for each node in the presented graph neural network. With the adoption of convolutional layers and dense layers as well as the similarity between each pair of students, the proposed approach can predict the future circumstance of the entire class. To evaluate the performance of our work, comparison experiments between several state-of-the-art algorithms and the proposed algorithm were conducted. The result from the experiments demonstrated that the graph neural network-based algorithm achieved competitive performance (sensitivity 91.24%, specificity 93.58%, and accuracy 89.79%) over the state-of-the-art.

1. Introduction

Due to the global pandemic novel coronavirus [1], most of the pedagogical practices in all countries have to be carried out online, which significantly impacts or even hampers the expansion of teaching and learning activities [2]. Therefore, it is necessary to ensure the students concentrate on the content of courses during the process of learning. Meanwhile, the teachers also need to comprehend the dynamic changes of the entire online class. Both requirements have rendered the automatic identification of the status of the students become a thorny task.

In recent decades, virtual reality (VR) related techniques [3, 4] have been extensively exploited in a plethora of applications [5] ranging from industrial manufacturing [6], healthcare [7], entertainment [8, 9], and education [10, 11]. Meanwhile, different types of Internet of things (IoT) have also been deployed in practice [12], e.g., commercial and industrial scenarios [13, 14], medical assistance [15], and smart cities [16]. Both of them have shown their capacity and potential value in multidisciplinary tasks. However, online education especially music teaching remains a frontier domain that needs to be developed. Meanwhile, it requires an effective mechanism for measuring the students' academic performance since the adoption of the current evaluating manner might not be advisable.

Meanwhile, the deep learning-based models that emerged in the last years have been widely accepted as a powerful instrument in a plethora of domains and applications including but not limited to computer vision [17, 18], natural language processing [19], data mining [20], computer-aided diagnosis [21], recommendation system [22], and forecasting [23–26]. A variety of architectures were proposed in deep learning orientation, including the convolutional neural network (CNN) [27], recurrent neural network (RNN) [28], autoencoder [29], and generative adversarial network (GAN) [30]. It is worth noting that the deep learning-based algorithms focus on unveiling the inner

patterns hidden in Euclidean samples and they usually neglect the non-Euclidean data with complicated associations and interdependency, e.g., social networks. Therefore, these deep models suffer from applying the common computational operators used in the CNN, RNN, autoencoder, and GAN to the graph domain.

On the other hand, to cope with the non-Euclidean data, numerous graph neural networks (GNNs) were put forward with different patterns and have shown their satisfactory performance in the recent period. As the work of Goodfellow et al. [30] is taken as an early work of GNN, the recently emerged GNNs can be roughly divided into four categories, including convolutional GNNs [31, 32], recurrent GNNs [33, 34], graph autoencoders (GAEs) [35, 36], and spatial-temporal GNNs [37, 38].

Although, plenty of GNN-based algorithms have been presented to cope with various machine learning tasks, e.g., handwritten signature recognition [31, 37], document discrimination [32, 35], ranking [33], program verification [34], and human activity detection [38]. In recent 5 years, for instance, Zhang et al. [39], proposed a deep graph clustering framework. First, a feature transformation module is built up of both the node and graph topology. Second, a graph embedding and a self-supervised learning mechanism are introduced to constrain graph embedding by using the graph similarity and self-learning loss. To deal with the molecular graph generation issue for drug discovery, Shi et al. [40] presented a flow-based autoregressive model, which combines both the autoregressive and flow-based approaches. Zheng et al. [41] proposed a graph multiattention network to predict traffic conditions for temporal phases ahead at differentiated positions within a traffic network graph. Aiming at addressing the fact-checking problem in the text, Zhong et al. [42] proposed a method for reasoning about the semantic level structure of evidence by using semantic role labeling. These algorithms have shown their performance in various scenarios. However, most of the GNN-based methods ignore feature information for each node in the graph.

Bearing the abovementioned analysis in mind, this work fills a gap in the literature by introducing VR integrated with IoT into musical pedagogy. To the best of our knowledge, this is also an early work for the employment of VR and IoT in online education. First of all, the proposed music teaching platform exploits the VR techniques to offer a uniquely immersive service for each student to experience an actual scenario without any hefty cost. For instance, the students can virtually perform in front of a large audience without a real stage. Second, it can collect the real-time status of the students by using IoT sensors. Accordingly, the students' physiological status including the heart rate, respiratory rate, temperature, and facial appearance can be captured. Third, the GNN model can then be introduced to represent the integral circumstance of the students to form the whole feature vector space of the online class. The personal information of all students captured from the IoT sensors is denoted as the nodes in a graph. The subjective judgment of the students' performance (positive or negative) made by the teachers is used as the ground truth in the training samples.

Afterwards, a binary-class (positive and negative) graph neural network is trained over the manually labeled data samples. Accordingly, the subtle variation in this graph can be identified and predicted with high accuracy. The presented GNNs are then tested on the input samples and the practical condition of the students can be generated by updating the features embedded in the graph nodes. In general, the proposed network is supposed to alleviate the workload of the teacher in an online music class, while the network's output can unveil the status of each student revealed by the VR and IoT devices.

To evaluate the performance of the proposed framework, we conducted comparison experiments between state-of-the-art GNN models on the practical data samples collected from an online music class. Accordingly, the experimental results output from the proposed GNNs indicated that the GNN-based pipeline can come up with the practical engineering requirements.

In general, the contributions of this work consist of the following:

(1) To our best knowledge, this is an early work for using VR and IoT in an online music learning platform

(2) A semisupervised learning framework is proposed to recognize the students' status with GNNs. The situation of the entire class can be revealed prominently.

(3) The experimental results demonstrate the effectiveness and efficiency of the proposed framework

2. Materials and Methods

2.1. Online Music Learning Framework and Dataset.
An online music learning system (as shown in the left component of Figure 1) was built upon VR integrated with IoT sensors. This system consists of head-mounted VR display devices (Skyworth S801) and IoT sensors. The VR apparatus has $100°$ field of view and an optic lens. Before each class started, the students were instructed to wear both the VR apparatus and the wearable IoT sensors.

For each student, the data samples (heart rate, respiratory rate, temperature, and facial appearance) were collected and came into a feature vector at every 5 minutes. Then, three teachers labeled each sample as positive or negative separately by using a majority voting mechanism. In total, 6,104 recordings (3,020 positives and 3,084 negatives) from 16 students were collected in this study.

2.2. Graph Neural Network

2.2.1. Problem definition.
It is supposed that n subjects (i.e., the students as input samples) are available, denoted by $S = [S_1, S_2, ..., S_n]$. Each student then can be denoted by one matrix $Si \in R^{m \times ni}$, where m represents the total number of IoT sensors mounted on each student, n is the length of the feature derived from the corresponding IoT sensor, and $n_i \in \{0, 1\}$. The weighted graph used in the proposed GNN is denoted as a tuple $G = (V, E, W)$, where V is the set of m vertices, E represents the set of edges, and $W \in R^{m \times ni}$ is the
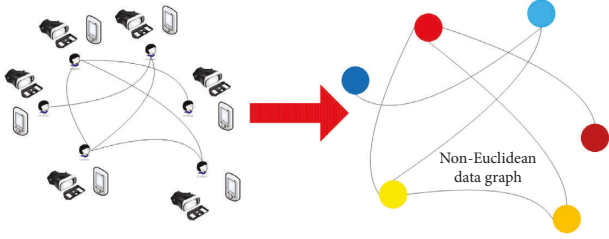
FIGURE 1: The pipeline from the online music learning platform with VR and IoT sensors to the non-Euclidean data structure.

adjacency matrix of the graph. Meanwhile, $W_{i,j}$ represents the weight of the edge from $v_i \in V$ to $v_j \in V$. In general, a global threshold is employed in the adjacency matrix to eliminate the irrelevant entries in the adjacent matrix or nodes in the graph. The space complexity of the proposed graph network is $O(n^2)$.

In general, the personal information of all students captured from the IoT sensors is denoted as the nodes in an undirected graph. Accordingly, the student's heart rate, respiratory rate, temperature, and facial appearance are integrated as a feature vector to represent the integral circumstance of the students. Meanwhile, the similarity between each pair of students is taken as the value of the edge between the two students denoted by two nodes. The edges with values lower than 0.4, which is set by conducting preliminary experiments, are eliminated from the graph.

In the proposed GNN model, the graph convolutional operator used can be considered as the spectral multiplication, which is equal to the convolution operator used in the temporal domain. Accordingly, the corresponding spectral filter can be realized by introducing the eigenfunctions of the normalized Laplacian into GNN.

$$L = I_m - D^{-1/2}WD^{-1/2}, \tag{1}$$

where D denotes the degree of the matrix and I_m is an identity matrix. Furthermore, the Laplacian matrix can be implemented by using Chebyshev polynomials [43]:

$$T_k(L) = 2LT_{k-1}(L) - T_{k-2}(L), \tag{2}$$

where $T_0(L) = 1$ and $T_1(L) = L$.

The K-ordered polynomial then yields unbiased K filters. Accordingly, the filtered outcome of the signal by K filters can be formulated as

$$o = g_\theta(L) * c = \sum_k 0_k \theta_k T_k(\bar{L})c, \tag{3}$$

where c denotes one IoT sensor mounted on a student, $L = 2/\lambda_{max}L - I_d$, and λmax is the maximal eigenvalue of the normalized Laplacian L. The output of the lth layer can then be expressed as

$$o_s^l = \sum_{i=1}^{F_{in}} g_{\theta_i^l}(L)c_{s,i}^l, \tag{4}$$

where F_{out} and F_{in} denote the output filter and input filter, respectively. $\theta_i^l \in R_K$ is the Chebyshev coefficient and $x_{s,i}^l$ is the input graph at layer l for student s.

A pooling layer is located at the end of the proposed GNN. In a fashion of semi-unsupervised learning, the feature map generated from the presented GNN can yield a pairwise association between the subjects. The intact GNN with the pooling operator is demonstrated in Figure 2.

2.2.2. Network Architecture. As described in Table 1, there are 3 pairs of convolutional layers integrated with the rectified linear unit (ReLU) as well as a pooling layer within the proposed GNN architecture. To guarantee the invariant scale for the graphs, the pooling is only incorporated at the end of the convolution operations. The dropout rate of every convolutional layer is 0.5.

In the initial stage, the training rate is set to 0.001 with fixed 500 iterations. Once the decrease in validation accuracy lasts for two consecutive iterations, then the learning rate is multiplied by 0.5. During the training process, 80% of the samples were used as the training set and the remaining 20% were taken as the testing set, and 100 of the training samples were used as the validation set. In total, there were less than 1,000 parameters in the proposed graph network. We trained the whole model by leveraging a back-propagating strategy and no overfitting was observed during training. Furthermore, the expected outcome of the proposed network is the binary status of each student (node in the graph) that is positive or negative.

3. Results and Discussion

3.1. Performance Evaluation Metrics. In this work, we used accuracy, sensitivity, and specificity in the experiments to measure the performance of the comparing methods.

(1) Sensitivity: the ratio between the number of true positives (TP, the samples are labeled as positive by the teachers; meanwhile, the method generates the correct result) and the number of all of the samples.

$$\text{Sensitivity} = \frac{TP}{TP + FN}, \tag{5}$$

where FN denotes the false negative (the outcome of one sample labeled as positive is negative).

(2) Specificity: the ratio between the number of true negatives (TN, the outcome of one sample labeled as negative is negative) and the number of all of the samples.

$$\text{Specificity} = \frac{TN}{TN + FP}, \tag{6}$$

where FP denotes the false positives (the outcome of one sample labeled as negative is positive).

(3) Accuracy: the ratio between the number of correctly identified subjects and the number of all of the samples.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP}. \tag{7}$$

Figure 2: The structure of the proposed GNN.

Figure 3: The optimal threshold value set in the proposed GNN.

Second, we conducted the experiments on the 6,104 samples captured from 16 students by using the proposed approach. For each student, the corresponding number of samples ranged from 197 to 405. It is widely accepted that the quantity and quality of the training samples significantly relate to the performance of machine learning models, as well as the deep learning models.

Learning-based algorithms: accordingly, the sensitivity, specificity, and accuracy yielded from our model can rise or fall for different students in an online class, as shown in Figure 4.

As shown in Figure 4, the accuracy for all of the students is greater than 80% by using the proposed approach, while the accuracy of students 5, 12, and 13 is almost 100%. Moreover, we also examined the samples collected from student 6 whose accuracy is the lowest among the students. We found that student 6 contains the fewest samples (197) compared with the other students.

Generally, the proposed method can provide satisfactory sensitivity and specificity for most of the subjects. However, its performance relates to the quantity and quality of the input samples. By improving the data sample, the proposed GNN can be a potentially valuable instrument for online learning performance evaluation in practice.

Third, we conducted the comparison experiments between the state-of-the-art and the proposed framework on the entire dataset that was manually collected. Experimental results as shown in Figure 5 demonstrate that the current work achieved competing performance over the state-of-the-art techniques. Although the sensitivity produced by [37] is better than the proposed approach, this work outperformed the state-of-the-art methods [31, 37, 38] in both the specificity and accuracy of the whole data samples.

4. Discussion

The results of the experiments show that the proposed method can provide a favorable outcome for the issue that has an intrinsic graph structure, and it can be useful for other cases rendering similar characteristics. However, there are some limitations of the presented approach that need to be addressed before it is applied to those tasks. For instance, it does not take temporal information into consideration. Therefore, the relations between sequential samples might be

Table 1: Details of the proposed GNN architecture.

Layer	Conv	ReLU	Conv	ReLU	Conv	ReLU	Pooling	Class
Channels	16	N/A	32	N/A	64	N/A	N/A	2
K-order	9	N/A	9	N/A	9	N/A	N/A	N/A
Stride	1	N/A	1	N/A	1	N/A	N/A	N/A

3.2. Experimental Results. In this work, the proposed GNN was implemented with the TensorFlow 2.0 deep learning architecture with Python as the programming language.

The edge between each pair of students in the graph would affect the whole structure of the graph. A global threshold is expected to determine if one edge is preserved or not. Thus, we tested the effects of various threshold values to determine the optimal threshold. In the preliminary test, only 9 different numbers were leveraged (0.1, 0.2, ..., 0.9). Meanwhile, to decrease the computation complexity of calculating the optimal threshold value, only the subsamples were used during this stage. According to the outcome from the subsamples in Figure 3, 0.4 is chosen as the threshold value in the following experiments.
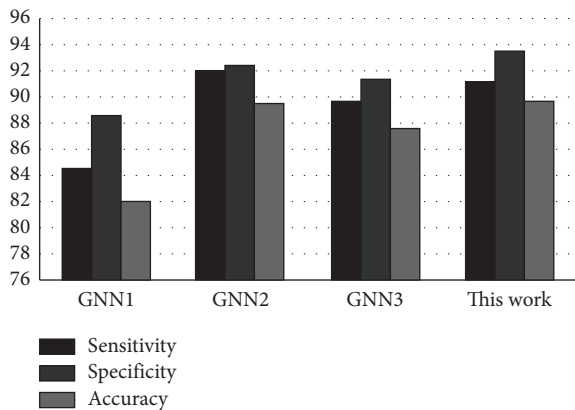
By leveraging it, personal information can be collected and formed into a non-Euclidean graph. Since it is difficult to employ the manners commonly adopted in an offline classroom, the identification and prediction of the students' real-time state deserve in-depth research for the online learning scenes. Bearing this in mind, we proposed a spatial-temporal GNN-based framework. Both the interdependent associations between the students and the corresponding development process can be unveiled from the presented GNN model.

To evaluate the performance of the proposed framework, comparison experiments were conducted between the state-of-the-art techniques and the proposed method. Experimental results demonstrated that the combination of VR and IoT, as well as GNN, should be taken as a potentially valuable instrument for online music learning.

In the future, the application of global average pooling (GAP) needs to be studied since it is supposed to decrease the number of parameters and eliminate overfitting and we will continue to delve into the online learning platforms and the applications of various machine learning-based algorithms into them.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Figure 4: The experimental outcome (sensitivity (%), specificity (%), and accuracy (%)) of the proposed approach for each student.

Figure 5: The comparison between the state-of-the-art and our work in terms of sensitivity (%), specificity (%), and accuracy (%).

neglected by using our approach. Moreover, this method is unable to handle less structured representations.

5. Conclusions

In this work, the VR apparatus and IoT sensors were introduced to implement an online music learning platform.

References

[1] L. Zhang, J. Chen, Q. Peng, and X. Shen, "Emergency management and protection measures during epidemics of novel coronavirus pneumonia," *Journal of Clinical Medicine in Practice*, vol. 24, 2020.

[2] C. L. Chang and M. Fang, "E-learning and online instructions of higher education during the 2019 novel coronavirus diseases (covid-19) epidemic," *Journal of Physics: Conference Series*, vol. 1574, no. 1, Article ID 012166, 2020.

[3] J. Steuer, "Defining virtual reality: dimensions determining telepresence," *Journal of Communication*, vol. 42, no. 4, pp. 73–93, 1992.

[4] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. J. Julier, and B. Macintyre, "Recent advances in augmented reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34–47, 2001.

[5] B. Schlake and M. Narayanan, "Virtual reality applications," in *Proceedings of the WESCON/94. 'Idea/Microelectronics'. Conference Record*, New york, NY, U.S.A, January 1994.

[6] T. Mujber, T. Szecsi, and M. Hashmi, "Virtual reality applications in manufacturing process simulation," *Journal of Materials Processing Technology*, vol. 155-156, pp. 1834–1838, 2004.

[7] A. A. Rizzo, M. Schultheis, K. A. Kerns, and C. Mateer, "Analysis of assets for virtual reality applications in neuropsychology," *Neuropsychological Rehabilitation*, vol. 14, no. 1-2, pp. 207–239, 2004.

[8] G. B. Newby, "Virtual reality and the entertainment industry. (cover story)," *Bulletin of the American Society for Information Science and Technology*, vol. 21, 1994.

[9] J. Lugrin, M. Cavazza, M. Palmer, and S. Crooks, "Artificial intelligence-mediated interaction in virtual reality art," *IEEE Intelligent Systems*, vol. 21, no. 5, pp. 54–62, 2006.

[10] A. Rienow, C. Lindner, T. Dedring, H. Hodam, and C. Jrgens, "Augmented reality and virtual reality applications based on satellite-borne and iss-borne remote sensing data for school lessons," *Photogrammetrie, Fernerkundung, GeoInformation*, vol. 88, 2020.

[11] O. L. Chávez, L. F. Rodríguez, and J. O. G. Gutierrez, "A comparative case study of 2d, 3d and immersive-virtual-reality applications for healthcare education," *International Journal of Medical Informatics*, vol. 141, 2020.

[12] J. Ni, K. Zhang, X. Lin, and X. S. Shen, "Securing fog computing for internet of things applications: challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 601–628, 2018.

[13] I. Culic, A. Radovici, and C. Rusu, "Commercial and industrial internet of things applications with the raspberry pi," *Prototyping IoT Solutions*, Apress, New york, NY, U.S.A, 2020.

[14] I. Nappi and G. de Campos Ribeiro, "Internet of things technology applications in the workplace environment: a critical review," *Journal of Corporate Real Estate*, vol. 22, no. 1, pp. 71–90, 2020.

[15] S. Y. Y. Tun, S. Madanian, and F. Mirza, "Internet of things (iot) applications for elderly care: a reflflective review," *Aging Clinical and Experimental Research*, vol. 33, pp. 1–13, 2020.

[16] A. S. Ibrahim, K. Y. Youssef, H. Kamel, and M. Abouelatta, "Traffic modelling of smart city internet of things architecture," *IET Communications*, vol. 14, no. 8, pp. 1275–1284, 2020.

[17] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, "A transfer-learning approach for accelerated mri using deep neural networks," *Magnetic Resonance in Medicine*, vol. 84, no. 2, pp. 663–685, 2020.

[18] A. Sezer and A. Altan, "Detection of solder paste defects with an optimization based deep learning model using image processing techniques," *Soldering & Surface Mount Technology*, vol. 33, 2021.

[19] M. A. Ayyoub, A. Nuseir, K. Alsmearat, Y. Jararweh, and B. Gupta, "Deep learning for Arabic nlp: a survey," *Journal of Computational Science*, vol. 26, pp. 522–531, 2017.

[20] G. Nguyen, S. Dlugolinsky, M. Bobak et al., "Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey," *Artifiicial Intelligence Review*, vol. 52, 2019.

[21] H. I. Suk, S. W. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis," *NeuroImage*, vol. 101, pp. 569–582, 2014.

[22] Q. Zhou, J. Wu, and L. Duan, "Recommendation attack detection based on deep learning," *Journal of Information Security and Applications*, vol. 52, Article ID 102493, 2020.

[23] A. Altan, S. Karasu, and S. Bekiros, "Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques," *Chaos, Solitons & Fractals*, vol. 126, pp. 325–336, 2019.

[24] S. Karasu, A. Altan, S. Bekiros, and W. Ahmad, "A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series," *Energy*, vol. 212, Article ID 118750, 2020.

[25] A. Altan, S. Karasu, and E. Zio, "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer," *Applied Soft Computing*, vol. 100, Article ID 106996, 2021.

[26] S. Karasu and A. Altan, "Crude oil time series prediction model based on lstm network with chaotic henry gas solubility optimization," *Energy*, vol. 242, Article ID 122964, 2022.

[27] A. Li, M. Yuan, C. Zheng, and X. Li, "Speech enhancement using progressive learning-based convolutional recurrent neural network," *Applied Acoustics*, vol. 166, Article ID 107347, 2020.

[28] A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Canada, May 2013.

[29] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[30] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial networks," *Advances in Neural Information Processing Systems*, vol. 3, pp. 2672–2680, 2014.

[31] J. Bruna, W. Zaremba, A. Szlam, and Y. Lecun, "Spectral networks and locally connected networks on graphs," 2013, https://arxiv.org/abs/1312.6203.

[32] M. Henaff, J. Bruna, and Y. LeCun, "Deep convolutional networks on graph-structured data," 2015, https://arxiv.org/abs/1506.05163.

[33] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 11, pp. 61–80, 2009.

[34] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, "Gated graph sequence neural networks," 2015, https://arxiv.org/abs/1511.05493.

[35] S. Cao, W. Lu, and Q. Xu, *Deep neural networks for learning graph representations*AAAI, Menlo Park, CA, U.S.A, 2016.

[36] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, pp. 1225–1234, ACM Press, San Francisco, CA, USA, May 2016.

[37] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, "Structured sequence modeling with graph convolutional recurrent networks," in *ICONIP (1), Ser. Lecture Notes in Computer Science*, L. Cheng, A. C. S. Leung, and S. Ozawa, Eds., vol. 11301, pp. 362–373, Springer, Berlin, Germany, 2018.

[38] A. Jain, A. R. Zamir, S. Savarese, and A. Saxena, "Structural rnn: deep learning on spatio-temporal graphs," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5308–5317, Las Vegas, NV, USA, June 2016.

[39] X. Zhang, J. Mu, H. Liu, and X. Zhang, "Graphnet: graph clustering with deep neural networks," in *Proceedings of the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Canada, June 2021.

[40] C. Shi, M. Xu, Z. Zhu, W. Zhang, M. Zhang, and J. Tang, "Graphaf: a flflow-based autoregressive model for molecular graph generation," 2020, https://arxiv.org/abs/2001.09382.

[41] C. Zheng, X. Fan, C. Wang, and J. Qi, "GMAN: a graph multi attention network for traffic prediction," 2019, https://arxiv.org/abs/1911.08415.

[42] W. Zhong, J. Xu, D. Tang et al., "Reasoning over semantic-level graph for fact checking," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Stroudsburg, PA, U.S.A, July 2019.

[43] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral fifiltering," 2016, https://arxiv.org/abs/1606.09375.

Research Article

A Deep Ranking Weighted Multihashing Recommender System for Item Recommendation

Suresh Kumar [ID],[1] **Jyoti Prakash Singh** [ID],[1] **Vinay Kumar Jain** [ID],[2] **and Avinab Marahatta** [ID][3]

[1]*Department of Computer Science and Engineering, NIT Patna, India*
[2]*Department of Management (PG), MIT World Peace University, Pune, India*
[3]*Center for Multidisciplinary Studies and Innovation (CeMuSI), Kathmandu, Nepal*

Correspondence should be addressed to Avinab Marahatta; avinab.marahatta@gmail.com

Collaborative filtering (CF) techniques are used in recommender systems to provide users with specialised recommendations on social websites and in e-commerce. But they suffer from sparsity and cold start problems (CSP) and fail to interpret why they recommend a new item. A novel deep ranking weighted multihash recommender (DRWMR) system is designed to suppress sparsity and CSP. The proposed DRWMR system contains two stages: the neighbours' formation and recommendation phases. Initially, the data is fed to the deep convolutional neural network (CNN). The significant features are extracted from CNN. The CNN contains an additional layer; the hash code is generated by minimising pairwise ranking loss and classification loss. Therefore, a weight is assigned to different hash tables and hash bits for a recommendation. Then, the similarity between users is obtained based on the weighted hammering distance; the similarity between users helps to form the neighbourhood for the active user. Finally, the rating for unknown items can be obtained by taking the weighted average rating of the neighbourhood, and a list of the top n items can be produced. The effectiveness and accuracy of the proposed DRWMR system are tested on the MovieLens 100 K dataset and compared with the existing methods. Based on the evaluation results, the proposed DRWMR system gives precision (0.16), the root mean squared error (RMSE) of 0.73 and the recall (0.08), the mean absolute error (MAE) of 0.57, and the $F-1$ measure (0.101).

1. Introduction

Recommendation systems (RS) have recently become common on numerous websites, recommending movies, e-commerce, music, and television programs [1]. Based on the information provided by the user, the RS recommends items for purchase [2]. Several RS have been introduced to predict the behaviour of users and provide better recommendations [3, 4]. In general, RS recommends, based on the individual interests of the user and their previous usage history, finding items with the highest preference [5]. RS is a great machine learning system to increase product sales [6, 7]. Recommendation helps the user to speed up the search process and makes it simple for them to obtain content that is interesting to them, as well as provide them with offers they would not have searched for [8, 9]. Furthermore, companies may attract customers by showing movies and TV shows relevant to their profiles [10].

The approaches to recommendation may be categorised as collaborative filtering (CF), content-based (CB), and hybrid based on the type of data gathered and how it is used in the RS [11]. CB filtering is frequently used in the RS design, which uses items' content to select general characteristics and qualities that suit the user profiles [12]. Then, users compare items to previously liked items and recommend the best matching-items. The CF system works based on a user-item relationship. Hybrid systems utilise both forms of information to prevent issues that arise when only one type is used.

Although the recommendation system has demonstrated its usefulness in various fields, it faces some problems: sparsity and cold start problems (CSP). As the name implies, a sparsity problem develops when customers do not rate an item while purchasing online, resulting in a sparsity of available ratings. The CF method goes through this problem because it uses a rating matrix. A CSP arises when a new product is added to the recommendation system, and no past ratings for that product are available. As a result, an RS must provide options for new users, which makes the accuracy of the recommendations low. Several research groups are working to develop an efficient and well-organised RS algorithm. The methods like local sensitive hashing (LSH) [13], Bayesian personalized ranking (BPR) [14], Cofi Rank-maximum margin matrix factorization (CRMF) [15], etc. are used for removing sparsity and CSP. However, the existing methods fail to generate deep hashing by minimising classification loss, ranking pairwise, and high storage costs, motivating us to develop an efficient method for recommending items.

1.1. Contribution. The following are the research work's contributions:

   (i) The hash code is generated by reducing the classification and pairwise ranking loss, which can handle the users with cold start and sparsity problems.

   (ii) A weight is introduced to hash tables and hash bits according to their performance.

   (iii) The weighted Hamming distance is used to determine user similarity.

   (iv) Bit diversity and similarity preservation are utilised to calculate the hash bit's weight, and the MAP score is used to evaluate the weight-based table-wise.

   (v) The proposed DRWMR system performs better for the item recommendation.

The following is a summary of the paper's structure: Section 2 provides an overview of existing works. Our proposed DRWMR System is explained in Section 3. In Section 4, the proposed method is subjected to experimental analysis. Finally, the paper concludes in Section 5.

2. Literature Review

Da'u et al. [16] proposed a deep learning method that uses Aspect-based opinion mining (ABOM) in a recommendation system. This method contains two portions. They are rating predictions and ABOM. In the first portion, we utilised multichannel deep convolutional neural networks for extracting aspects. In the second portion, we add aspect-based ratings into a machine's tensor factorisation to predict overall ratings. This method increases the accuracy of the recommendation system, but the time consumption is more significant.

Ye and Liu [17] introduced Collaborative Topic Regression (CTR) and three novel granulation methods for the recommendation strategy. The three granulation methods are LDA-based, PMF-based, and CTR-based, based on building granular structures and conducting granulation. LDA-based and PMF-based methods are developed to extract granular features from content and feedback information. The CTR-based method is designed to be joined with LDA- and PFM-based methods to produce multilevel recommendation information and interpretable granular features. These methods were introduced to overcome the problem of time cost and decision cost. However, this method does not consider users' dynamic preferences.

Chen et al. [18] offered a Dynamic Decay CF (DDCF) recommendation method based on the user's interest. This method has four stages: clustering of items, identification of the interesting level, specification of the decay function, and preference prediction. In item clustering, similar items are grouped without any predefined parameters. Then, in the second stage, based on the number of ratings and time records in the cluster, we can identify the interesting level of users. In the third stage, decay functions describe the preference evolution at every level. Finally, the similarities between the users are calculated based on the decay rates, and future preferences are predicted. This method accurately predicts the user's interests, but the time consumption is high.

Beg et al. [19] introduced a chaotic based reversible data transform method (CRDTM) for preserving privacy in data mining for the recommendation system. This method will dynamically produce the values of RDT parameters during processing time, but it does not need the parameters during recovery. This method can substitute for the standard RDT algorithm, in which memory and bandwidth are considered significant factors. The RDT algorithm is used in the mobile app recommendation area. This method decreases the execution time, but accuracy is low.

Abbasi-Moud et al. [20] offered a context-aware recommendation system for tourism. It consists of three steps. 'Users' preferences are extracted from their reviews and comments in the first step. In the second step, tourist 'attractions' characteristics are extracted from the tourist reviews. In the third step, recommendations are given based on the 'user's preference and its similarity with the tourist attractions' characteristics and contextual information. The contextual information utilised in this method includes location, weather, user preferences, and time. This method gives high accuracy in tourism recommendation. However, the weather may change according to the season, so it's challenging to make a recommendation.

Zhang et al. [21] introduced a novel hybrid probabilistic matrix factorization method for distinguishing between items' attractiveness and users' preferences for a recommendation. It consists of two sub-divisions. One division attempts to predict the rating scores of users by extracting the user's personal preferences from auxiliary data. Another division attempts to model the textual interest of items for different users. These two sub-divisions are formed into a unified framework using a global objective function. This method gives high accuracy.

Choe et al. [22] offered a hierarchical model based on a recurrent neural network (HMRNN) for considering users' item usage histories in time series and sequences in a recommendation. This method contains two layers: a long-term (LT) and a short-term (ST). The ST sequences are handled by the ST layer, typically found in older ones. The LT layer recollects the data from the preceding ST sequences and distributes it to the lasting time. This method increases the amount of stored user history. However, it does not take the consideration of temporal properties.

3. DRWMR System

We propose a novel deep ranking weighted multihashing recommender (DRWMR) system to suppress sparsity and CSP by considering CSP and sparsity issues. This method has two stages: the neighbours' formation and recommendation phases. Initially, the user-item rating data of the user is fed into the neighbours' formation stage, and deep convolutional neural networks are used to extract text features. The deep neural network consists of an additional layer (hash table). The hash codes are produced by reducing the classification loss and pairwise ranking loss. The similarity preservation is preserved by pairwise ranking loss, and the prediction error is minimised by classification loss. Because each hash bit performs differently in the RS, it is not easy to treat them all equally. Therefore, a weight is assigned to different hash tables and hash bits for a recommendation. Then, the similarity between users is obtained based on the weighted Hamming distance—the similarity between users helps to get the neighbourhood formation for the active user. Finally, the rating for unknown items can be obtained by taking the weighted average rating of the neighbourhood and producing a list of the top n items. Figure 1 illustrates the architecture of the proposed DRWMR system.

3.1. Neighborhood Formation Phase.
The user-item rating data is given to the deep CNN, which is used to extract the features. Then, a hash code is generated by minimising classification loss and ranking pairwise loss. Therefore, a weight is assigned to different hash tables and hash bits. According to weighted Hamming distance, the most similar users are identified as active users.

3.1.1. Deep CNN.
The convolution and max pooling layers are used to extract the features. The demographic data of the user is fed to the convolution pooling layer. The convolution pooling layer is formed with the rectified linear activation function in the first three layers. The max pooling process is utilised in the initial convolution pooling layers, and the average pooling process is used in the final convolution pooling layers. They mostly use the original raw data to extract high-level representations. In the first three convolutional layers, we use 32, 32, and 64 with stride 1 for every convolution layer and a kernel size of 5×5. 3×3 is the size for pooling operations, while 2 is the stride for each pooling layer. The fully-connected layer uses a rectified linear activation function and a dropout layer with a 0.5 dropout ratio.

A tangent-like activation is used in the hash code layer to output hash codes. The softmax function is utilised as the activation function in the classification layers to preserve semantic similarity. The classification and hash code layers are sequentially learned in every training epoch. In the first fully-connected layer, the unit number is 500, the hash code's length is equal to the second fully-connected layer's output, and the number of classes is similar to the third fully-connected layer's output. A deep CNN architecture is constructed with this framework to learn the function of nonlinear transformation $\alpha(*)$ using the input as data. The hash bit is determined by using

$$a = \text{sign}(\alpha(K), \tag{1}$$

where K denotes the raw data, and sign (y) signifies the sign function.

The deep neural network consists of an additional layer (hash table). The hash codes are produced by reducing the classification loss and pairwise ranking loss. The similarity preservation is preserved by pairwise ranking loss, and the prediction error is minimised by classification loss. For a pair of users, the pairwise-loss function can be defined as follows:

$$S(a_i, a_j, I_{i,j}) = \frac{1}{2}(1 - I_{i,j})H_d(a_i, a_j) + \frac{1}{2}I_{i,j}\max(l - H_d(a_i, a_j), 0). \tag{2}$$

Such that $a_i, a_j \in \{+1, -1\}^h$,

where a denotes the user K's binary hash code, $H_d(a_i, a_j)$ signifies the Hamming distance between a_i, and a_j. The similarity between a_i and a_j is denoted by $I_{i,j}$, it is defined as

$$I_{i,j} = \begin{cases} 0 \text{ if } a_i, \text{ and } a_j \text{ are similar} \\ 1 \text{ if } a_i, \text{ and } a_j \text{ are dissimilar}. \end{cases} \tag{3}$$

The Hamming distance between the input user e_i, and its pairwise user is calculated to produce an arranged hamming list. The pairwise ranking loss for an input user e_i is defined in the following equation:.

$$P(a_i) = \sum_{j=1}^{M_i} P(S(a_i, a_j, I_{i,j})) = \sum_{j=1}^{M_i} S^2(a_i, a_j, I_{i,j}), \tag{4}$$

where M_i denotes the pairwise user's number for user e_i. Users similar to e_i should have short Hamming distances, therefore they seem at the top of the arranged Hamming list, whereas dissimilar data seem at the bottom. On the other hand, data at the improper positions of the Hamming list of e_i has substantial loss values. We aim to decrease the overall loss function for a training database with M user, which is given in the following equation:

$$P = \sum_{i=1}^{M} \sum_{j=1}^{M_i} P(a_i, a_j). \tag{5}$$

Such that $a_i, a_j \in \{+1, -1\}^h, i = 1, \dots \dots M$.

Since the hash codes are in binary, the fitness function is nondifferentiable. As a result, using the gradient descent approach is challenging for optimising the fitness function.
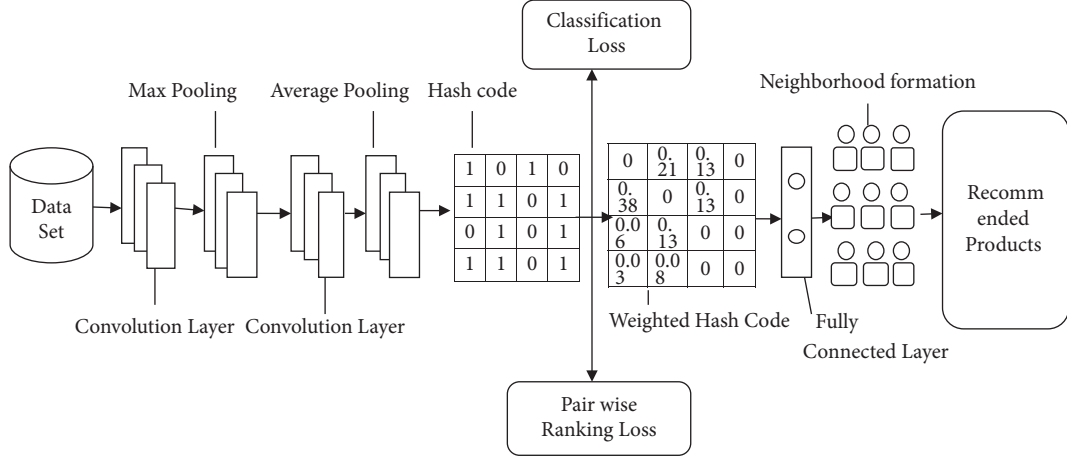
FIGURE 1: Architecture of the proposed DRWMR system.

Therefore, to overcome this problem, a tanh approximation function is used instead of the sign function. The tanh-like function can be used to estimate a_i's hash code, which is given as follows:

$$p_i(a_i) = \frac{e^{a_i} - e^{-a_i}}{e^{a_i} + e^{-a_i}}. \tag{6}$$

The Euclidean distance $E_d(p_i,)PJ$ between two users, can be further estimated as the Hamming distance $H(a_i, a_j)$ using the calculated hash codes determined in the following equation:

$$E_d(p_i, p_j) = \left\| a_i - a_j \right\|_2^2. \tag{7}$$

A regularisation term is included to reduce the quantisation loss. The final loss function is determined in the following equation:

$$P = \sum_{i=1}^{M} \sum_{j=1}^{M_i} \frac{1}{2} \left\{ \left(I_{i,j} E_d(p_i, p_j) + \frac{1}{2} I_{i,j} \max \left(l - E_d(p_i, p_j), 0 \right) \right)^2 \right\} + \sum_{i=1}^{m} \alpha \left\| p_i - a_i \right\|_2^2. \tag{8}$$

Such that $a_i, a_j \in \{+1, -1\}^h, i = 1, \ldots \ldots M,$

where α denotes the regularisation term's parameter coefficient. The classification layer outputs the recommendation for the category. As a recommendation function, the softmax function is used. The classification loss is determined in the following equation:

$$C_i = -\sum_{j=1}^{o} \hat{D}_{ij} \frac{e^{o_{X_{ij}}}}{\sum_{k=1}^{o} e^{o_{X_{ik}}}}. \tag{9}$$

Where o and o_x_{ij} signify the class's number and the recommended output of j's class for user e_i.

3.1.2. Weighted Multihashing.

A weighted multihash code is used based on the loss in the hamming distance. It multiplies the bitwise weight and table-wise weight to minimise loss. Each hash bit performs differently in the recommendation task, so it is not fair to treat them all the same. Therefore, a weight is assigned to different hash tables and hash bits. Bit diversity and similarity preservation are integrated for the bitwise weighting. The MAP value is utilised for each hash table to compute the weight-based table-wise. The weight-based table-wise and bitwise may change for dissimilar users. This method adjusts the two weights, and their

product determines the final weight. The bit b in table r final weight is defined as follows:.

$$WT_{b,r} = WT_{\text{bit}_b} \times WT_{\text{bit}_r}. \tag{10}$$

Similarity preservation (P) measures a hash bit's semantic similarity. The hash bit a_r''s similarity preservation (P) is calculated in the following equation:

$$Q(a_b) = \frac{1}{2} (a_b' I a_b - U), \tag{11}$$

where U signifies the training sample's number, and a_b' signifies the a_b's transpose.

The bit diversity is utilised to calculate the hash bit's performance. The difference between each user's hash bit is crucial for maintaining recommendation efficiency. Therefore, every hash bit must be independent. The correlation between two hash bits is used to determine bit diversity, which can be determined by the above formula.

Initially, a bitwise correlation metric $R = \left\{ r_{Q_{ij}} \mid i, j, \ldots \ldots h \right\} \in Z^{h \times h}$ can be generated by calculating $r_{Q_{ij}}$ for bit a_i and a_j pair. The bitwise correlation is measured using the Kappa value, which is determined in the following equation:

$$r_q = \frac{\varnothing_1 - \varnothing_2}{1 - \varnothing_2} \in [-1, 1], \tag{12}$$

where $\varnothing_1$, and $\varnothing_2$ denote the hash bit's probabilities. Eqn can determine the $_1$, and $\varnothing_2$ probabilities.

$$\varnothing_1 = \frac{w + z}{V}$$
$$\varnothing_2 = \frac{(w + x)(w + y) + (y + z)(x + z),}{v^2} \tag{13}$$

where w, x, y, and z denote the nonnegative variables indicating the sample's number that fulfills the relevant row and column requirements. The bit-correlated matrix $M = \{m_1, \ldots \ldots, m_h\} \in Z^{1 \times h}$ is formed by assigning the correlated coefficient for each bit. Finally, the bita_r's bit diversity weight is calculated by the following equation:

$$K(a_b) = \frac{1 - f_b}{2}, \tag{14}$$

where f_b signifies the bth bit's correlated coefficient. To generate a final bitwise weight WT_{bit_b}, the terms mentioned above are first adjusted and multiplied, which is determined in the following equation:

$$WT_{\text{bit}_b} = Q(a_b) \times K(a_b). \tag{15}$$

To combine multiple hash tables, table-wise weight (WT_{bit_r})for each hash table is determined using the mean average precision. It is specified in the following equation:

$$A = \frac{1}{P} \sum_{b=1}^{n} \frac{b}{P_b} \times \text{rel}_b, \tag{16}$$

where n signifies the size of the dataset, P denotes the similar user's number in the dataset that relates to the checking user, P_b denotes the similar user's number at the top of the dataset.

3.2. Recommendation Phase.
The neighbours obtained from the previous phase are used to generate a recommendation. This similarity between the active user and its neighbours is utilised to forecast the final rating for an unknown rating of item i. In this regard, the ultimate anticipated rating provided by user u for any item i is calculated as the following:

$$\text{Pre d}_{u,i}^U = \overline{r}_u + \frac{\sum_{u' \in N(u)} \text{sim}(u, u') \times (r_{u',i} - \overline{r}_{u'})}{\sum_{u' \in N(u)} \text{sim}(u, u')}, \tag{17}$$

where $N(u)$ is the user's neighbour, u and $\text{sim}(u, u')$ denote the similarity between two users.

The prediction value of each item can be sorted in descending order, and the top n items from this sorted list of items can be recommended to the user.

The similarity between the users can be obtained by the following equation:

$$\text{sim}(u, u') = \frac{\sum_{i \in I} (r_{u,i} - \overline{r}_u)(r_{u',i} - \overline{r}_{u'})}{\sqrt{\sum_{i \in I} (r_{u,i} - \overline{r}_u)^2} \sqrt{\sum_{i \in I} (r_{u',i} - \overline{r}_{u'})^2}} \tag{18}$$

where I is the set of items,$r_{u,i}$ rating of given to item $i$by user u, $\overline{r}_u$ average rating of user u.

4. Experimental Results and Analysis

This section contains the experimental setup and a description of the datasets for comparative analysis.

4.1. Dataset Description.
The movie Lens 100 K dataset contains numerous users' demographic data. There are 1682 movies in the dataset, with a total of 10,0000 ratings from 943 users. For this evaluation, data were gathered from the https://www.kaggle.com/prajitdatta/movielens-100k-dataset. It consists of the user's age, ID, occupation, and items provided. The datasets are used for 80% of the training, while the remaining 20% is used for testing.

4.2. Simulation Setup.
The proposed DRWMR system is implemented in python; the initial learning rate is 0.001, and after 1000 iterations, it lowers exponentially by 0.04. The batch size is 200, and the parameter coefficient is 0.01.

4.3. Evaluation Criteria.
The model's performance is evaluated using six metrics: recall, precision, F1-score, MAE, RMSE, and prediction accuracy.

(i) Recall:

It is the ratio of the total number of relevant recommendations to the actual or true number of relevant recommendations for a new user.

$$\text{RL} = \frac{F_r}{T_r} \tag{19}$$

where F_r denotes the recommendation related to the new user and T_r signifies the real amount of related recommendations.

(ii) Precision:

This metric indicates the precision of the process, i.e., whether the generated recommendations are appropriate for new users.

$$\text{PN} = \frac{F_r}{\text{tot}_r} \tag{20}$$

where F_r denotes the recommendation related to the new user and tot_r signifies the recommended item's total quantity.
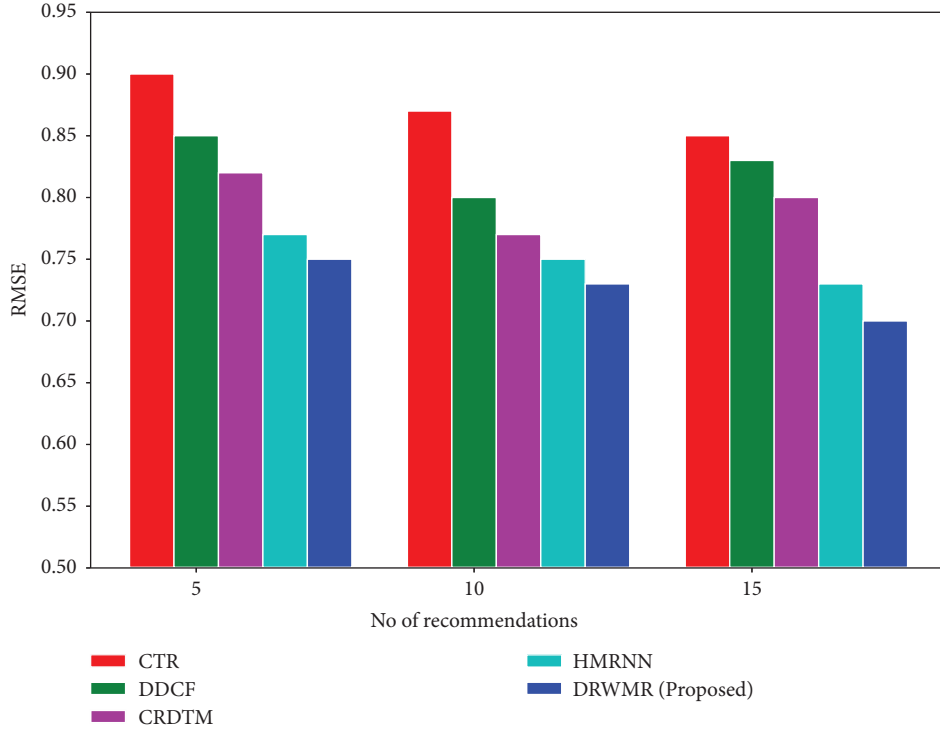
FIGURE 2: RMSE analysis.

(iii) F1-score:

This score shows the experiment's accuracy based on the recall and precision measures and is calculated by the following equation:

$$F = \frac{2 \times RL \times PN}{PN + RL} \qquad (21)$$

(iv) RMSE:

It is measured by determining the difference between the predicted and observed values, which is given in the following equation:

$$R = \sqrt{\frac{1}{Q} \sum_{j=1}^{Q} \left(k_j - \hat{k}_j \right)^2}. \qquad (22)$$

(v) MAE:

The magnitude of the difference between the expected and observed values is computed to determine it, which is given in the following equation:

$$M = \frac{1}{Q} \sum_{j=1}^{Q} \left| k_j - \widehat{K_j} \right|, \qquad (23)$$

where Q signifies the user's total amount, $\widehat{K_j}$KJ denotes the predicted values, and KJ signifies the observed value.

4.4. Comparative Analysis with Existing Methods. The proposed DRWMR system is tested on the movie Lens 100 K dataset, and metrics such as recall, precision, RMSE, MAE, and F1-score are compared to existing approaches such as CTR [17], DDCF [18], CRDTM [19], and HMRNN [22].

4.4.1. Quantitative Analysis. Figure 2 illustrates the RMSE analysis of our proposed DRWMR system with existing methods such as CTR, DDCF, CRDTM, and HMRNN. Our research introduces a novel loss function that minimises the classification loss and ranks pairwise losses. It produces hash codes with high recommendation accuracy and more similar information. Therefore, our proposed DRWMR system reduces the error compared to the existing methods for the top 5, 10, and 15 recommendations. The proposed DRWMR system (0.73) has a low RMSE when compared with existing techniques such as CTR (0.87), DDCF (0.80), CRDTM (0.77), and HMRNN (0.75) for the top 10 recommendations.

Figure 3 shows the MAE analysis. The smallest Hamming distance between the users results in the recommendation to the new user. The weighted Hamming distance is used to determine how similar the users are. So, the errors can be reduced in the recommendation system. The MAE is low for the proposed DRWMR system (0.57) when compared with existing methods such as CTR (0.75), DDCF (0.72), CRDTM (0.705), and HMRNN (0.653) for the top 10 recommendations.

Figures 4 and 5 show the recall and precision analyses. In the proposed DRWMR system, a hash table is built as an additional layer. A table and bitwise weight method are presented based on the performance to attain a better recommendation performance. The proposed DRWMR system assigns weights to various hash bits and hash tables. Also, a hash code is generated based on minimising loss,
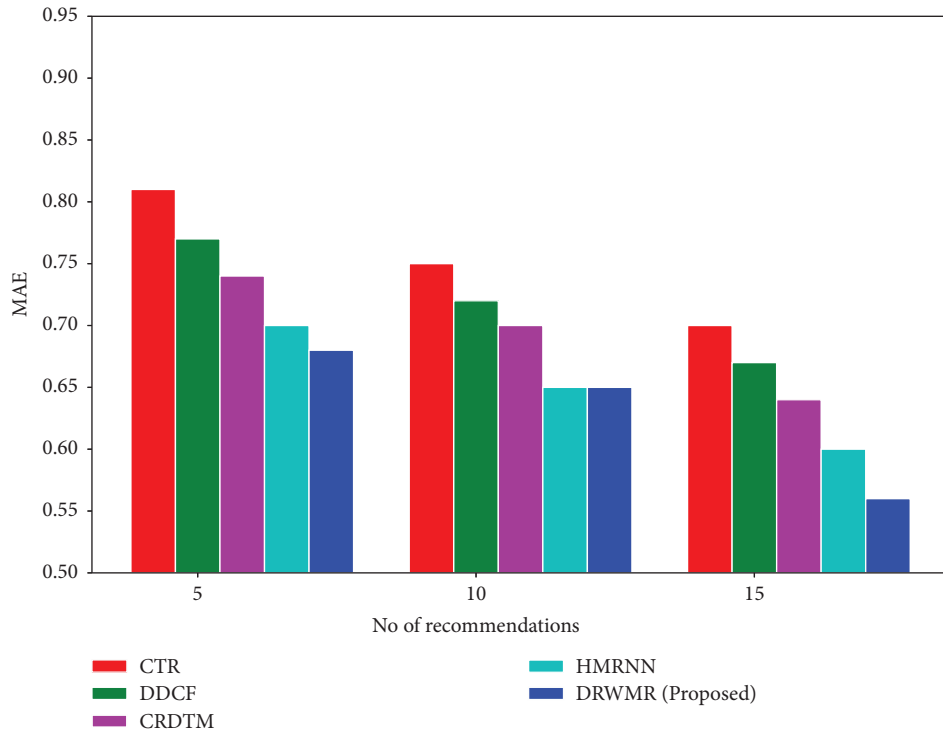
Figure 3: MAE analysis.
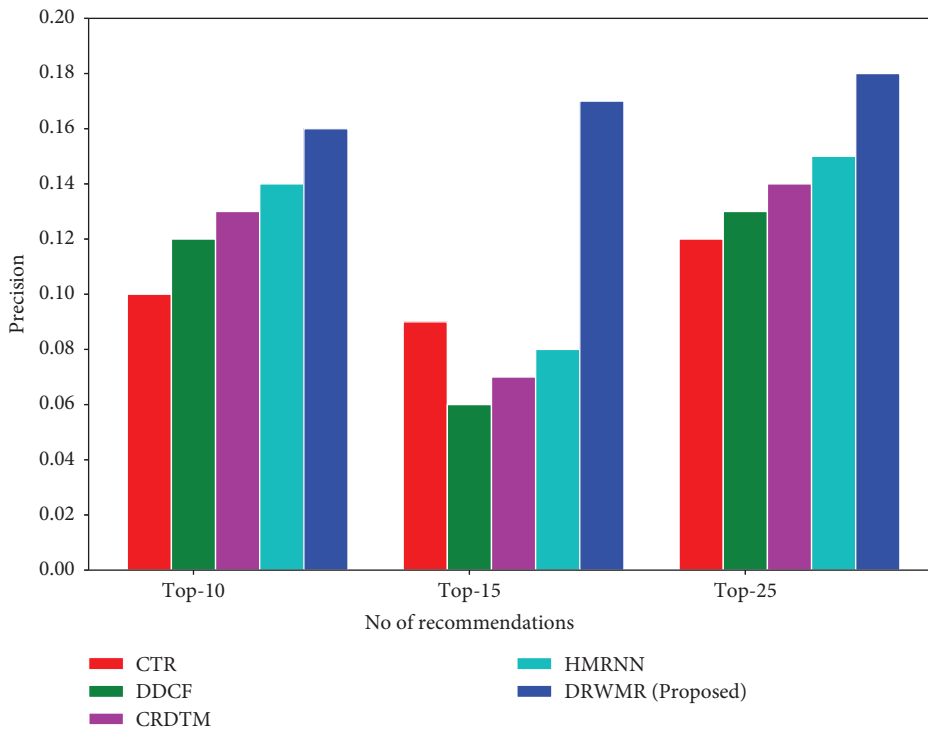
Figure 4: Precision analysis.

which can handle users with cold start and sparsity problems. This mechanism attains higher precision and recall. The precision is high for the proposed DRWMR system (0.16) when compared with existing methods such as CTR (0.10), DDCF (0.12), CRDTM (0.129), and HMRNN (0.14) for the top 10 recommendations. The recall is high for the proposed DRWMR system (0.08) when compared with existing methods such as CTR (0.04), DDCF (0.041), CRDTM (0.045), and HMRNN (0.06) for the top 10 recommendations.
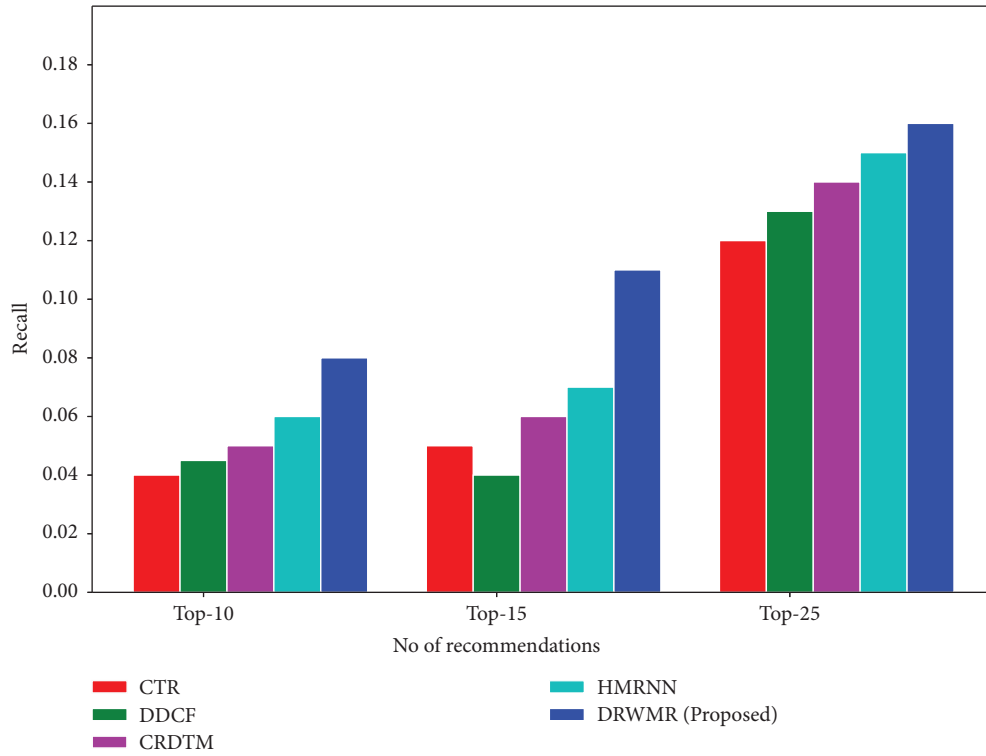
Figure 5: Recall analysis.
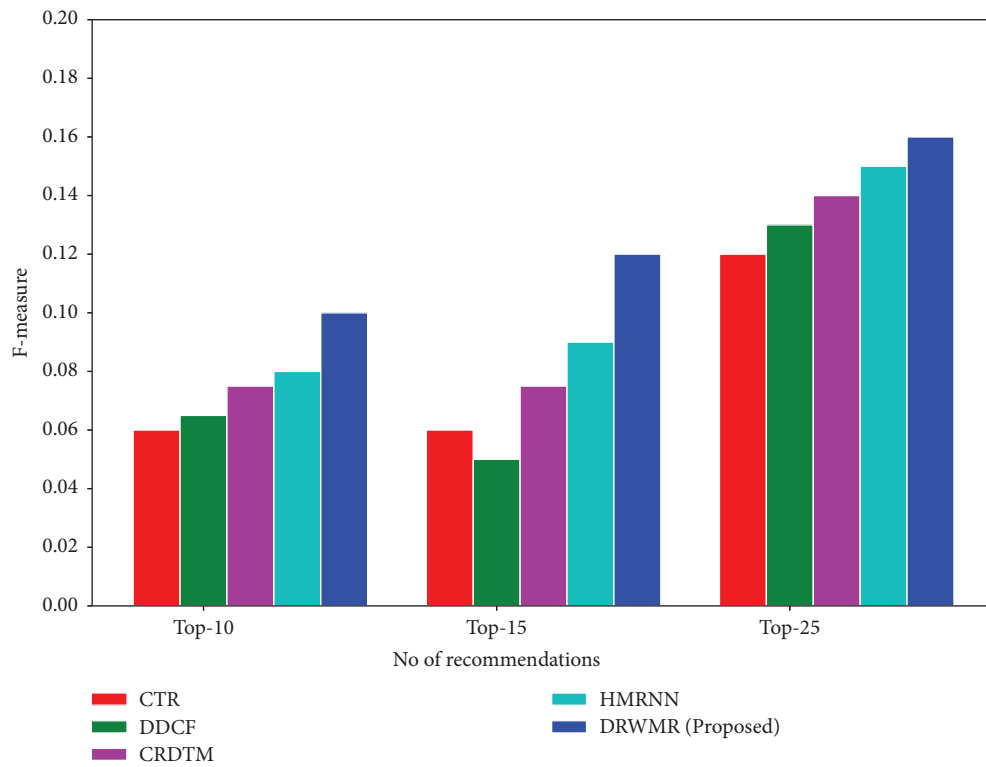
Figure 6: *F*-measure analysis.

Figure 6 shows the $F-1$ measure analysis. The F-measure is high for the proposed DRWMR system (0.101), when compared with existing methods such as CTR (0.06), DDCF (0.064), CRDTM (0.077), and HMRNN (0.08) for the top 10 recommendations.

5. Conclusion

The RS has become vital to social networking and business apps such as Flipkart, Amazon, YouTube, and others. Therefore, we introduced the DRWMR system for an accurate recommendation initially. User-item rating data is fed into CNN to extract important features to reduce data sparsity. Then, the hash code is generated by minimising pairwise rank loss and classification loss. A weight is assigned to different hash tables and hash bits. According to weighted Hamming distance, the most similar users were obtained. Finally, the rating of unknown items can be obtained using the weighted average rating of similar users and active users. As a result, we have seen that our proposed DRWMR system performs better in recommending items. The proposed DRWMR system is quantitatively measured by precision (0.16), recall (0.08), RMSE (0.73), MAE (0.57), and F-measure (0.101). It shows better performance in the recommendation system.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative filtering and deep learning based recommendation system for cold start items," *Expert Systems with Applications*, vol. 69, pp. 29–39, 2017.

[2] M. Kawasaki and T. Hasuike, "A recommendation system by collaborative filtering including information and characteristics on users and items," in *Proceedings of the In 2017 IEEE Symposium Series on Computational Intelligence (Ssci)*, Orlando, FL, USA, 2017 November.

[3] D. Das, L. Sahoo, and S. Datta, "A survey on recommendation system," *International Journal of Computer Application*, vol. 160, no. 7, pp. 6–10, 2017.

[4] P. Lopes and B. Roy, "Dynamic recommendation system using web usage mining for e-commerce users," *Procedia Computer Science*, vol. 45, pp. 60–69, 2015.

[5] S. Pandya, J. Shah, N. Joshi, H. Ghayvat, S. C. Mukhopadhyay, and M. H. Yap, "A novel hybrid based recommendation system based on clustering and association mining," in *Proceedings of the In2016 10th International Conference on Sensing Technology (ICST)*, IEEE, Nanjing, China, 2016 November.

[6] J. P. Verma, B. Patel, and A. Patel, "Big data analysis: recommendation system with Hadoop framework," in *Proceedings of the 2015 IEEE International Conference on Computational Intelligence & Communication Technology*, Ghaziabad, UP, India, 2015 February.

[7] S. Shaikh, S. Rathi, and P. Janrao, "Recommendation system in E-commerce websites: a graph-based approached," in *Proceedings of the In 2017 IEEE 7th International Advance Computing Conference (IACC)*, Hyderabad, India, 2017 January.

[8] T. Rutkowski, J. Romanowski, P. Woldan, P. Staszewski, R. Nielek, and L. Rutkowski, "A content-based recommendation system using neuro-fuzzy approach," in *Proceedings of the In2018 IEEE International Conference on Fuzzy Systems*, pp. 1–8, Hyderabad, India, 2018 July.

[9] S. Parvatikar and B. Joshi, "Online book recommendation system by using collaborative filtering and association mining," in *Prcoceedings of the In 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–4, Trivandrum Kerala, 2015 December.

[10] W. T. Chu and Y. L. Tsai, "A hybrid recommendation system considering visual information for predicting favorite restaurants," *World Wide Web*, vol. 20, no. 6, pp. 1313–1331, 2017.

[11] G. Suganeshwari and S. S. Ibrahim, "A survey on collaborative filtering based recommendation system," in *Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC–16')*, Springer Cham, India, 2016 February.

[12] J. Jooa, S. Bangb, and G. Parka, "Implementation of a recommendation system using association rules and collaborative filtering," *Procedia Computer Science*, vol. 91, pp. 944–952, 2016.

[13] L. Qi, X. Wang, X. Xu, W. Dou, and S. Li, "Privacy-aware cross-platform service recommendation based on enhanced locality-sensitive hashing," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1145–1153, 2021.

[14] G. Li, T. Zhu, J. Hua et al., "Asking images: hybrid recommendation system for tourist spots by hierarchical sampling statistics and multimodal visual bayesian personalized ranking," *IEEE Access*, vol. 7, Article ID 1265, 2019.

[15] M. Weimer, A. Karatzoglou, Q. Le, and A. Smola, "Cofirank-maximum margin matrix factorisation for collaborative ranking Advances in Neural Information Processing Systems," in *Prceedings of the 21st Annual Conference on Neural Information Processing Systems*, Vancouver, British Columbia, Canada, 2007 December.

[16] A. Da'u, N. Salim, I. Rabiu, and A. Osman, "Recommendation system exploiting aspect-based opinion mining with deep learning method," *Information Sciences*, vol. 512, pp. 1279–1292, 2020.

[17] X. Ye and D. Liu, "An interpretable sequential three-way recommendation based on collaborative topic regression," *Expert Systems with Applications*, vol. 168, Article ID 114454, 2021.

[18] Y. C. Chen, L. Hui, and T. Thaipisutikul, "A collaborative filtering recommendation system with dynamic time decay," *The Journal of Supercomputing*, vol. 77, no. 1, pp. 244–262, 2021.

[19] S. Beg, A. Anjum, M. Ahmed et al., "Dynamic parameters based reversible data transform (RDT) algorithm in recommendation system," *IEEE Access*, vol. 9, pp. 110011–110025, 2021.

[20] Z. Abbasi-Moud, H. Vahdat-Nejad, and J. Sadri, "Tourism recommendation system based on semantic clustering and sentiment analysis," *Expert Systems with Applications*, vol. 167, Article ID 114324, 2021.

[21] X. Zhang, H. Liu, X. Chen, J. Zhong, and D. Wang, "A novel hybrid deep recommendation system to differentiate user's preference and item's attractiveness," *Information Sciences*, vol. 519, pp. 306–316, 2020.

[22] B. Choe, T. Kang, and K. Jung, "Recommendation system with hierarchical recurrent neural network for long-term time series," *IEEE Access*, vol. 9, pp. 72033–72039, 2021.

Retraction

Retracted: Recognition of Continuous Music Segments Based on the Phase Space Reconstruction Method

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Huang and Y. Hu, "Recognition of Continuous Music Segments Based on the Phase Space Reconstruction Method," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4099505, 11 pages, 2022.

Research Article

Recognition of Continuous Music Segments Based on the Phase Space Reconstruction Method

Xuesheng Huang [ID] [1] **and YanQing Hu** [ID] [2]

[1] *School of Music and Dance, Quanzhou Normal University, Quanzhou, Fujian 362000, China*
[2] *Dean's Office, Quanzhou Normal University, Quanzhou, Fujian 362000, China*

Correspondence should be addressed to Xuesheng Huang; hxs@qztc.edu.cn

Piano score recognition is one of the important research contents in the field of music information retrieval, and it plays an important role in information processing. In order to reduce the influence of vocals on the progress of piano notes and restore the harmonic information corresponding to piano notes, the article models the harmonic information and vocal information corresponding to piano notes in the frequency spectrum. We use the phase space reconstruction method to extract the nonlinear feature parameters in the note audio and use some of the parameters as the training set to construct the support vector machine (SVM) classifier and the other part as the test set to test the recognition effect. Therefore, the method of adaptive signal decomposition and SVM is introduced into the signal preprocessing link, and the corresponding recognition process is established. In order to improve the performance of the support vector machine, the article uses measurement learning method to obtain the measurement learning and uses the measurement learning to replace the Euclidean distance of the Gaussian kernel function of the support vector machine. The SVM method of adaptive signal decomposition and the SVM method of principal component analysis are introduced into the preprocessing process of the note signal, and then the preprocessed signal is reconstructed in phase space, and the corresponding recognition process is established. The method of directly reconstructing the phase space of the original signal has higher accuracy and can be applied to the note recognition of continuous music segments. The final experimental results show that, compared with the current popular piano score recognition algorithm, the recognition accuracy of the proposed piano score recognition algorithm is improved by 3.5% to 12.2%.

1. Introduction

With the development of digital information technology in the 21st century, many applications based on music signal analysis, such as automatic labeling of music, music score translation, and music retrieval, have higher requirements for analysis results [1]. However, due to the complex composition of music signals, even for fully digitized signals, there is still no more complete method to analyze them in detail [2]. Piano score recognition is one of the important research issues in music signal processing. It plays an important role in song cover recognition, audio matching, and music recommendation systems [3]. From a signal perspective, we are concerned about two characteristics of audio: time and frequency. Most of the audio signal

processing is also focused on the time-frequency analysis of music and frequency images and various statistical methods cannot achieve a good trade-off between time and frequency, so that audio processing has not been greatly improved [4]. The design of piano score notes feature and the selection of measurement learning models have always been the most active research content in the field of piano score note recognition. The content of piano music notes is not only an expression of the creator's emotions, but also an external representation of the music spectrum structure [5]. Therefore, the recognition of musical piano notes is inseparable from the knowledge of various aspects such as signal processing. According to the different processing domains, the single-tone estimation method can be divided into four categories: time domain method, frequency domain method,

time-frequency domain method, and cepstral domain method [6]. Among them, the short-time autocorrelation method, the harmonic peak method, the wavelet analysis method, and the cepstrum method have outstanding detection effects in each category. However, occasionally there will be situations where higher harmonic components are greater than the fundamental frequency components in music signals, which will make the detection error of the harmonic peak method [7].

Liu et al. [8] first compressed the signal spectrum energy to 12 musical scales based on music harmonic information and music theory knowledge and proposed to use 12-dimensional Pitch Class Profile (PCP) as the note feature of piano music scores. The piano score notes of the music are transcribed. Corpuz [9] proposed the Harmonic Pitch Class Profile (HPCP) feature and used it in the piano score recognition system. The experimental results show that the HPCP feature can effectively reduce the influence of the type of instrument on the notes of the piano score. Mello et al. [10] use the Enhanced Pitch Class Profile (EPCP) feature combined with the harmonic product spectrum and the PCP feature. EPCP superimposes and maps the harmonic energy by multiplication instead of summation, so that the distinction between simple chords is more obvious, and the effect of relatively complex piano notation patterns is poor, so it is rarely used. In addition to that, in recent years, scholars have further introduced ideas such as matrix analysis into the design of musical notes in piano scores. For example, Sun et al. [11] proposed a new type of NNLS-PCP feature based on Non-Negative Least Square (NNLS) solution. Before mapping the spectrum to the PCP vector, this feature obtains the frequency values corresponding to 84 notes through a priori knowledge and generates the dictionary based on this. Each element in the dictionary represents the fundamental frequency and its harmonic frequency corresponding to a note. Then they used the approximate x of the obtained spectrum and performed the mapping between the approximate spectrum and the PCP feature [12]. Through prior knowledge, this method can effectively solve the crosstalk of the fundamental frequency and harmonic frequency of the note during the mapping [13]. In order to enhance the robustness of the PCP feature to the notes of piano scores played by different timbre instruments, Liang et al. [14], based on the discrete cosine transform and the Mel frequency, proposed a Mel logarithmic PCP feature. A more practical type of multitone estimation method is the cyclic extraction method, which was first proposed by Patel and Chauhan [15]. Its idea is to regard the multitone signal as the superposition of multiple single-tone signals and then use a two-step loop to perform one by one estimate. In theory, it can deal with a mixture of countless single tones, and the detection result is particularly good for a mixture composed of random frequencies. But for music signals, the detection results are not satisfactory [16, 17]. In response to the existing problems, Noroozi et al. [18] also proposed an improved loop extraction method based on the principle of spectral smoothing. In this way, it is possible to avoid other harmonics that repeat the fundamental frequency and the detected fundamental frequency

from being removed from the signal after the current detection, which will affect the next detection. However, there is no universal standard for the principle of spectral smoothing, which reduces its applicability [19–21].

This article uses the method of measurement learning, according to the characteristics of the notes of the piano score, with supervised learning from the prior knowledge of the question itself to a measurement learning measurement matrix. Through this distance measurement, the original feature space is projected to a space with higher class discrimination, so that, in the projected feature space, the distribution of feature vectors with the same label is more compact, and the interval between feature vectors with different labels is larger. We further use measure learning to improve the Euclidean distance in the original SVM Gaussian kernel function, so that the improved measure learning-based SVM has better category discrimination. This article focuses on the aspect of frequency and on the basis of obtaining its accurate information, to the greatest extent to approximate the original signal. Through the processing of the approximation signal, the purpose of analyzing the original signal is achieved, and combined with the statistical measurement learning method, the single-tone signal is recognized. By collecting audio information from a single frequency level to obtain an accurate approximation of the signal, the approximation format is unified, the complex time-frequency analysis is transformed into a signal to matrix transformation, and the approximation format is detected with the measurement learning method, which improves the adaptability.

2. Construction of Piano Music Note Recognition Algorithm Model Based on Measure Learning Support Vector Machine

2.1. Support Vector Machine Hierarchical Spatial Distribution. Support Vector Machine (SVM) is a maximum boundary classifier, which has good robustness to outliers. SVM maps the feature vectors one by one into labeled outputs through strict mapping relationships. SVM is established on the basis of the dimension theory and the principle of structural risk minimization in the statistical learning theory. It is based on the limited sample information (training sample) in the model complexity and the learning ability to find the best compromise, so that it has the best generalization ability [22–24]. Figure 1 shows the hierarchical spatial distribution of support vector machines.

The key step of feature acquisition is to model the harmonic components corresponding to the piano score notes in the music signal spectrum and the nonharmonic components corresponding to the sparse and large noises and to obtain an optimization problem through augmented Lagrangian multiplication. The submethod is used to solve the problem, and finally the characteristic with better performance is obtained. This feature can remove large and sparse noise in the signal and reconstruct the harmonic information in the music signal, so that the feature contains more stable and pure harmonic information [25–27].
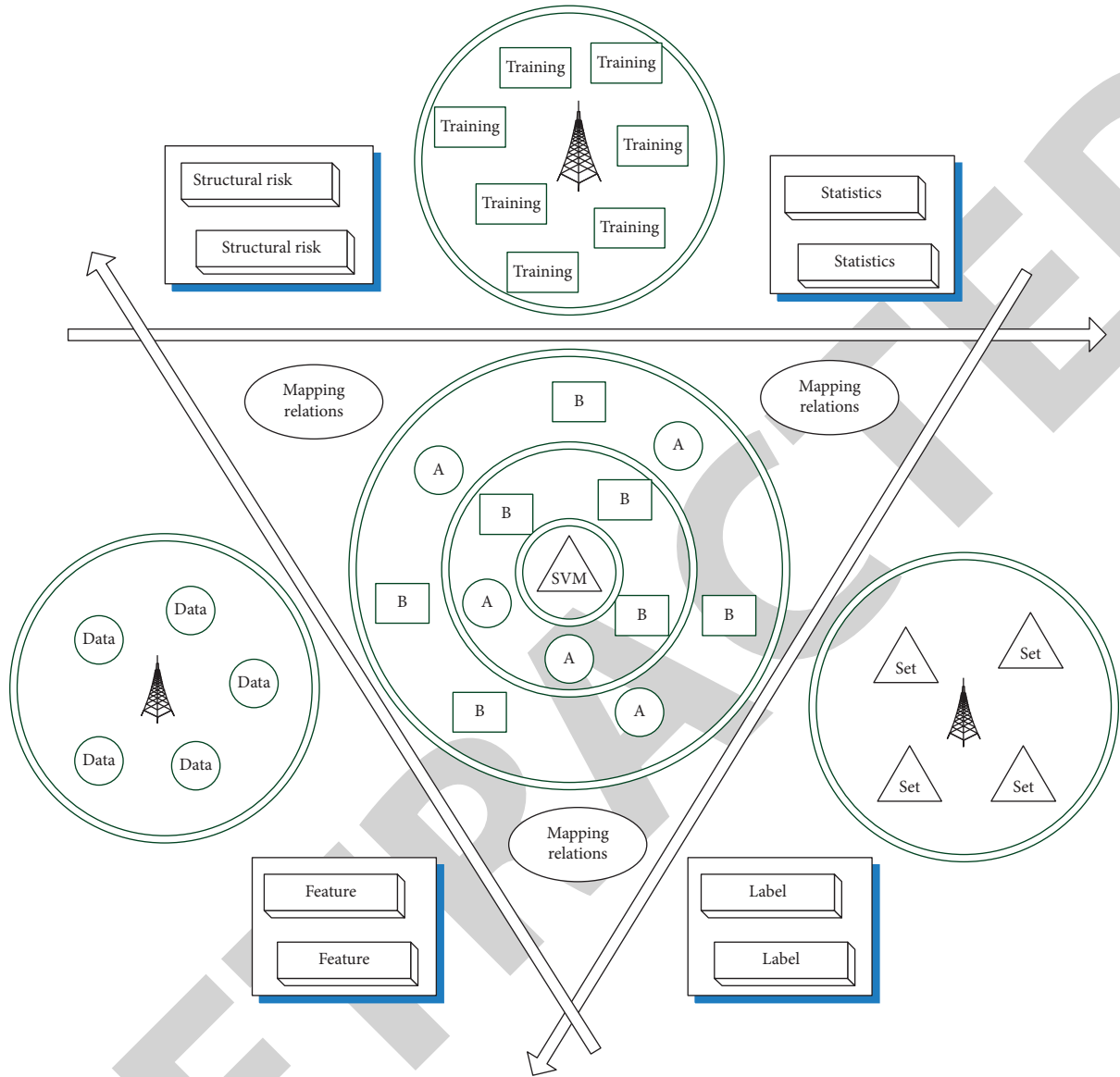
FIGURE 1: Hierarchical spatial distribution of support vector machines.

$$X[t|x(t)] = \{t \in R | x(1), x(2), \ldots, x(t)\},$$
$$y(n) = A \times X(n) - t \times x(n-1). \tag{1}$$

The signal preprocessing uses the empirical mode decomposition method. SVM is a smoothing process for a signal, and the result is to decompose the fluctuation or trend of different scales in the signal to produce a series of data series with different characteristic scales. Each series is called an intrinsic mode function.

$$h(x) = \begin{cases} \varepsilon(i,n) \times x(t), & 0 < t < 1, \\ \delta(i,n) \times x(t), & t > 1, \\ y(n) \times x(t), & t < 0. \end{cases} \tag{2}$$

After adding the signal preprocessing link, it is necessary to determine the number of characteristic subsignals after decomposition and determine the optimal phase space

reconstruction parameters corresponding to each characteristic subsignal. When the cumulative contribution rate of the current k-level principal components reaches 85%, it indicates that the previous k-level principal components basically contain all the information contained in the test data.

$$f(x+1) - f(x) = \partial \frac{y(x,n)}{\|d(n)\|^2} d(n). \tag{3}$$

In addition, because the feature vectors corresponding to some piano score types have a certain degree of similarity, the feature vectors corresponding to different piano score notes are likely to have a small Euclidean distance, which makes the classification accuracy of the support vector machine not high.

$$\overline{x^2}(k,n) - |x(k,n)|^2 - 2|\overline{x(k,n)}| \times \cos \alpha = 0. \tag{4}$$

We hope that the classification process is a process of measurement learning. In fact, there are many classifiers that meet the above requirements, and what we are looking for is an optimal classifier that satisfies the maximum separation of data points belonging to two different classes, so this hyperplane is called the maximum separation hyperplane.

$$\begin{cases} \dfrac{\alpha(x) - \alpha(x - 1)}{2\pi x} = 1, \\[2mm] \dfrac{\beta(x) - \beta(x - 1)}{2\pi x} = 1. \end{cases} \tag{5}$$

For a musical instrument, the sound it plays is of course caused by its vibration, but the strength and duration of the frequency vibration that constitutes the note must be measured by the amplitude.

$$g(x, k) = \frac{1}{n} \times \sum_{i=1}^{n} (1 - \alpha(x)) \times (1 - \beta(x)) \times x(t). \tag{6}$$

2.2. Measure Learning Recognition Algorithm.

In general, the adaptive signal decomposition of the music signal using the empirical mode decomposition method will generate a large number of subsignals, and most of the subsignals contain relatively little information. If all are determined as features in the case of subsignals, it will greatly increase the time required for the construction and recognition of the classifier and reduce the efficiency. Therefore, it is necessary to determine the signal containing more information as the characteristic subsignal from the decomposed subsignals.

$$W(x) = \begin{bmatrix} \dfrac{x(1)}{g(k,1)} & \dfrac{x(2)}{g(k,2)} & \cdots & \dfrac{x(n)}{g(k,n)} \\[2mm] 1 & 0 & \cdots & 0 \\[2mm] 0 & 1 & \cdots & 0 \\[2mm] 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{7}$$

By observing the decomposed subsignal image and its corresponding spectrum image, it can be found that the first two subsignals after decomposition contain more information. Therefore, the first two subsignals after decomposition are determined as characteristic subsignals. After determining the number of characteristic subsignals, it is also necessary to determine the delay time and embedding dimension corresponding to each characteristic subsignal during phase space reconstruction.

Therefore, we cannot measure the gap between the selected hypothetical model and the real model. The spectrum before compression is called the original spectrum; then the compressed spectrum is compared with the original spectrum. For example, if the sampling coefficient is set to 2, then the first vertex in the original spectrum corresponds to the second vertex in the compressed spectrum. It corresponds to the third vertex in the compressed spectrum; when the

original frequency spectrum and all compressed frequency spectrums are correspondingly multiplied, the result can clearly show the position of the fundamental frequency.

Figure 2 shows a schematic diagram of the measure learning recognition algorithm. After obtaining a classifier, we can approximate the error with the real model with certain quantities we have already mastered. The most intuitive method is to use sample data to classify the sample data with the obtained classifier and compare the result with the real classification result (still sample data). The difference between them is called the empirical risk. Algorithmically, the real number frame is used in the approximation process to avoid the solution of the high-order band matrix. The recognition process uses the normalized vector principle to standardize the amplitude range, so that the features at any time have a standard representation, which is more conducive to training and recognition. Theoretically speaking, an audio signal generated by a monophonic melody, the amplitude matrix A, obtained after the frame is approximated, each row of which should contain only the amplitude of one note, that is, the basis of the note.

Table 1 shows the combination of musical note audio signal attributes. For a single-tone melody, it is composed of different notes, and after being played by a specific instrument, each note contains the unique overtone combination of this instrument, that is, the frequency combination. We can determine to construct the vector with the amplitude of the frequency as the feature and input it into the SVM as a training sample. From the knowledge of music theory, we also know that the different combinations of fundamental frequency and overtones are precisely the characteristics of the sound of the musical instrument, and it is the mark that distinguishes it from other musical instruments.

In theory, an audio signal is generated by a monophonic melody (only one note played at any time). Therefore, if we know the frequency composition of each note played by a certain musical instrument, then for the calculated row vector in A, the combination of its nonzero elements in turn corresponds to a note. After obtaining, we count the number of consecutive notes, so that the note and note duration of this audio signal can be obtained.

2.3. Weight Optimization of Piano Score Data.

The method of principal component analysis SVM is a statistical method to reduce the dimensionality of the variable space. This method uses the orthogonal projection method to transform the original data from the high-dimensional variable space to the low-dimensional feature variable space, reducing the dimensionality of the variable space under the premise of the characteristic information, and the original data is retained. In this way, the analysis of the high-dimensional original data with the correlation between the variables is transformed into a low-dimensional space, thereby reducing the difficulty of analysis. In the case of unknown rules, an effective tool for classification is SVM.

Autocorrelation is effective for detecting repetitive patterns in a signal, such as detecting the period of a periodic
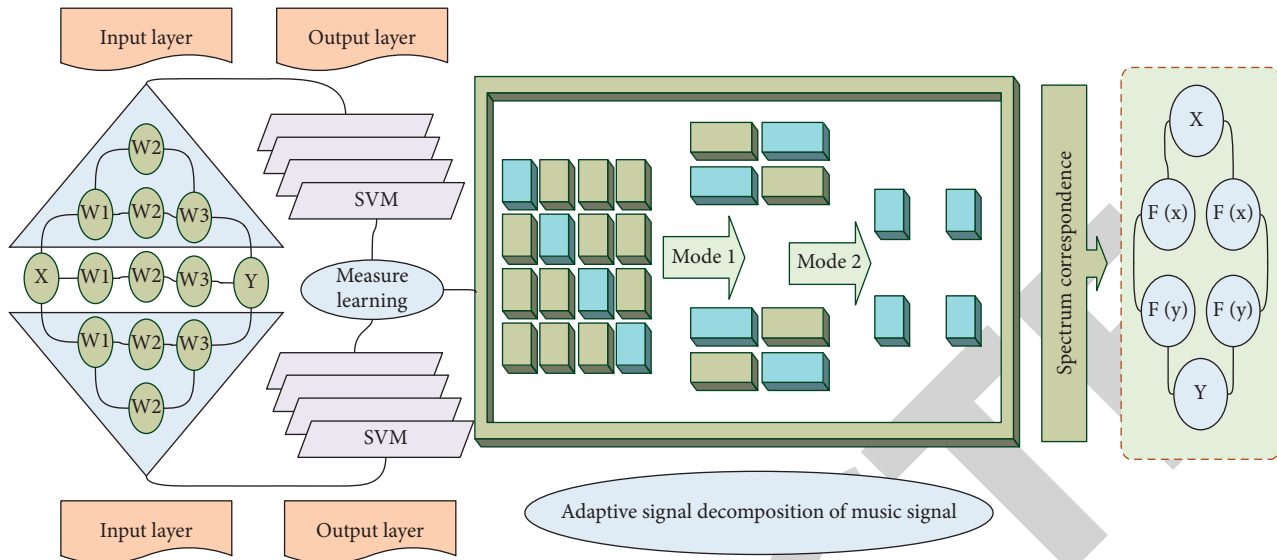
FIGURE 2: Schematic diagram of measure learning recognition algorithm.

TABLE 1: Musical note audio signal attribute combination.

Node serial	Frequency (Hz)	Duration ratio (%)	Standard deviation
1	2000	87	1.81
2	4000	91	0.98
3	6000	79	2.02
4	8000	56	1.96
5	10000	54	0.59
6	12000	69	0.64
7	14000	88	0.77

FIGURE 3: Comparison of broken lines of detection error for different groups of note data.

signal covered by noise. In addition, the results of SVM analysis on other kinds of notes also show that the first four eigenvalues can represent the main components in the nonlinear matrix. Figure 3 shows the comparison chart of the detection error broken lines of different groups of note data. It can be seen that, in the process of node length from 100 sampling points to 700 sampling points, the error ratio is also reduced from 3.6% to 1.5% in the worst case, but relative to the influence of low sampling times, the length of node is helpful to the accuracy of pitch detection and the accuracy of pitch extraction, but the effect is not as great as the impact of low sampling.

Since the timbre characteristics of each instrument, that is, the composition of the frequency corresponding to the notes, are unknown, we need to build training data and input it into the SVM to build a classification model. First, the music signal is distributed to 9 nonoverlapping bands (nonoverlapping bands) through the filter group. Figure 4 shows the feature vector processing process of musical note data information. In each frequency band, we use the autocorrelation algorithm to find the pitch of its corresponding frequency band. In the final stage, the pitches obtained in all frequency bands are fitted, and the desired pitch of the music is synthesized. Each row vector containing frequency and amplitude in A is exactly a feature vector containing audio

information corresponding to the note, so it can be input into SVM for training.

For the traditional short-term analysis method, if the window width is 50 ms, taking the wav file with the highest sampling accuracy as an example, only 2205 samples can be obtained. Both accuracy and breadth are far from meeting the requirements of audio signal analysis. If the window width is increased to 1 s, the frequency domain resolution can reach the minimum difference 1 for distinguishing music signals. For a song with a tempo of 200, the 32th note lasts only 37.5 ms, which obviously cannot be detected. This is the bottleneck of traditional short-time analysis, which is a trade-off between the resolution of the time domain and the frequency domain. Because the music itself has signal stability in a short period of time, it exhibits a certain period, so the autocorrelation method can be used to detect the period of the music, and of course the fundamental frequency of the signal can be detected. We trade off model complexity and learning ability based on small sample data, hoping to
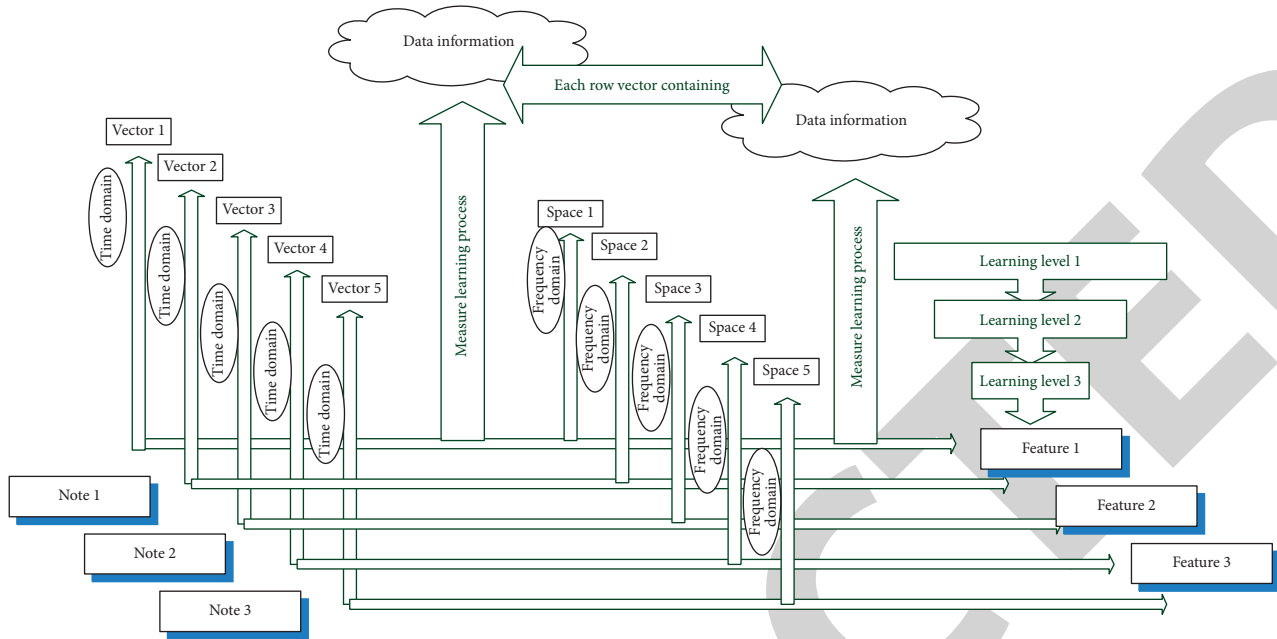
Figure 4: Feature vector processing process of musical note data information.

achieve the best generalization ability measurement learning method. It can be found that the cumulative contribution rate of the first four eigenvalues can reach more than 90%.

3. Results and Analysis

3.1. Feature Extraction of Musical Notation Data. The masking of notes and missing notes during the performance of the music will have a certain impact on the recognition and transcription of the final piano score. Therefore, the extraction of robust PCP features is divided into two steps: reconstruction of the harmonic information in the signal spectrum and scale mapping. The frame method starts from the whole audio signal, first obtains accurate frequency information from a large window, and uses the known frequency to construct an approximation of the original signal. In this way, the main frequency information of the original signal is kept as close as possible to its time domain information, so as to achieve the analysis of the music signal.

The two-tone connection recognition results of common piano notes are given; that is, the first note starts at 0 s, the second note starts at 0.5 s, and the number of recognition samples is 44100. From the data, it can be seen that, for the two-tone case, the classification is completely correct and does not include nonaudio internal notes; in most cases, the first note gets more recognition results, and the error is completely caused by this. This process is generally about 10–15 ms. Although theoretically speaking, the second note starts playing at 0.5 s, but its effective time will have a lag of 10–15 ms. That is to say, when the frequency contained in the second note is large enough to be classified, there is a 10–15 ms lag from the theoretical time.

Table 2 shows the statistical classification of the recognition results of piano score data. For the reliability and convenience of category labeling, the songs in the test music library and their corresponding labels are all from the music library on the Internet. We have selected a total of 480 songs in 6 categories. In this experiment, all music songs are in mp3 format, and their sampling rate is adjusted to 11025 Hz, and the bit rate is 16 kbps. It provides a software package for reading mp3 format music through MATLAB software. In this experiment, the length of 1024 is used for each frame. In the octave critical frequency band filter, because the sampling frequency of the test data is 11024 Hz, when it is decomposed into the critical frequency band, only the first 7 frequency bands are included in the effective frequency band, because the upper limit of the effective frequency is 5513 Hz. When calculating the 44th and 60th frames, the experimental data are as follows.

Each row of the amplitude matrix A obtained after the frame approximation should only contain the amplitude of one note. The usual method for detecting the starting point of a note is to find the "transient" area in the music signal; there are many definitions for the transient area, such as energy mutation points, short-term changes in the signal spectrum, etc. The above reduces the size of the data volume but refers to the audio signal as the input of a highly low sampling detection function and the role of the detection function to show the "instantaneous" in the original signal.

In this experiment, only the first 120 seconds of each song are intercepted for training or testing. Therefore, the time length of most songs is 3–5 minutes, so the first 120 seconds can already contain the basic characteristics of the song. For songs up to 120 seconds, a loop sampling method is adopted; that is, the beginning part of the song is connected to the end of the song. Figure 5 shows the fan-shaped distribution diagram of the recognition accuracy of different sampling units. In the 120 seconds of each song, 2.6 seconds are used as a music segment; each music segment is further divided into smaller unit music frames. The sampling

TABLE 2: Statistical classification of piano score data recognition results.

Node number	Frame number	Resolution (Hz)	Time delay (ms)	Error (%)
1	10	3.60	15	3.10
2	20	3.50	17	2.70
3	30	2.90	18	3.30
4	40	2.70	16	3.10
5	50	3.10	19	2.90
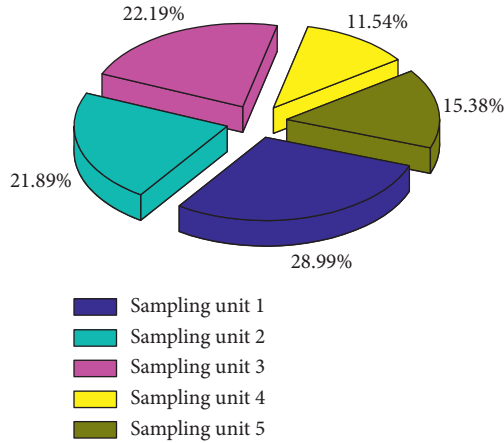
FIGURE 5: Fan-shaped distribution diagram of recognition accuracy of different sampling units.

frequency occupied by each time period is 28.99%, 21.89%, 22.19%, 11.54%, and 15.38%, respectively. Each music segment is further divided into 64 music frames. Each music frame contains 512 sampling points, and adjacent frames overlap 64 sampling points.

That is, the amplitude of the fundamental frequency and the overtone constituting the note and the amplitude of the frequencies contained in the notes that are not sounded at that moment are all zero. At the same time, due to the natural attenuation of audio, when the original feature data is processed by SVM, the projection feature of 95% of the original feature information is retained. After the SVM projection processing, the final feature of the dataset is reduced to 13 dimensions, and the feature of the dataset is reduced to 8 dimensions. Then, the SVM algorithm based on each kernel function is used to classify the data after SVM dimensionality reduction. It can be seen that after the original feature data is processed by SVM, the classification accuracy of each SVM is improved. This is because the original feature data after SVM projection and dimensionality reduction can effectively reduce aliasing and redundant information, thereby improving the accuracy of the final classification.

3.2. Simulation of Measure Learning Recognition Model. Firstly, the music fragment is sampled and processed with a sampling rate of 11025 Hz, and an audio signal with a length of node is generated. For the audio signal mode of this length, we use the window function $W(n)$ to preprocess the audio signal, and the window shift length is 50% of the window width to obtain the windowed waiting. Then, we

multiply the signal matrix by the Fourier transform matrix, the force is an $N \times N$ square matrix, and the rows of the matrix are the orthonormal basis, so that the frequency spectrum matrix of the framed signal is obtained as M. In the experiment, phase space reconstruction is performed on the original note signal, the first subsignal, and the second subsignal, the SVM principal component is extracted, and then the nonlinear feature construction is carried out. The piano score data set adopts MIDI format digital score files that contain piano score information such as pitch, beat, time, chord, speed, and channel.

In order to obtain more features related to the difficulty of piano scores, this paper adopts the related features of difficulty as the feature space of this paper. Therefore, the feature space in this article includes all the difficulty and semantic features except the fingering complexity (which belongs to the information of the score labeling level, and MIDI scores does not contain this information), key signature (key signature is meta tag information); except for the fingering feature (also belongs to the label layer feature) all the features, a total of 22 features, form a 22-dimensional feature vector to represent the MIDI score. Given the recognition results of common triplet within 1 s, 50 samples are selected for recognition from the approximation data. Figure 6 shows a matchstick graph of the attenuation rate of the note data signal based on measurement learning. From the figure, we can see that there are still notes that will get more detection results. This is caused by the crossover of notes, especially when the sound is not deliberately muted during performance, so that each note is naturally attenuated and played.

Either from the audio waveform graph or from the approximate amplitude graph, it can be clearly found that the previous note still has a strong amplitude at the note boundary. After normalization, the note with a weaker amplitude will be detected. However, although reducing the sampling rate increases the error percentage, the statistically obtained duration does not change much, and it has no effect on the translation of the score.

In the experiment, the 20 samples of each note signal were taken as test signals for the identification experiment. Figure 7 shows the SVM-based note data recognition deviation box diagram. In addition, considering many situations in actual piano learning and teaching, music scores are divided into 4 difficulty levels, and a large number of music websites generally provide digital music scores with 4 difficulty levels (easy, beginning, intermediate, and advanced). The scalability of the algorithm in this paper is suitable for practical applications. We also collected 400 MIDI scores from the large music website to form a data set with 4 difficulty levels, each with 100 MIDI scores.
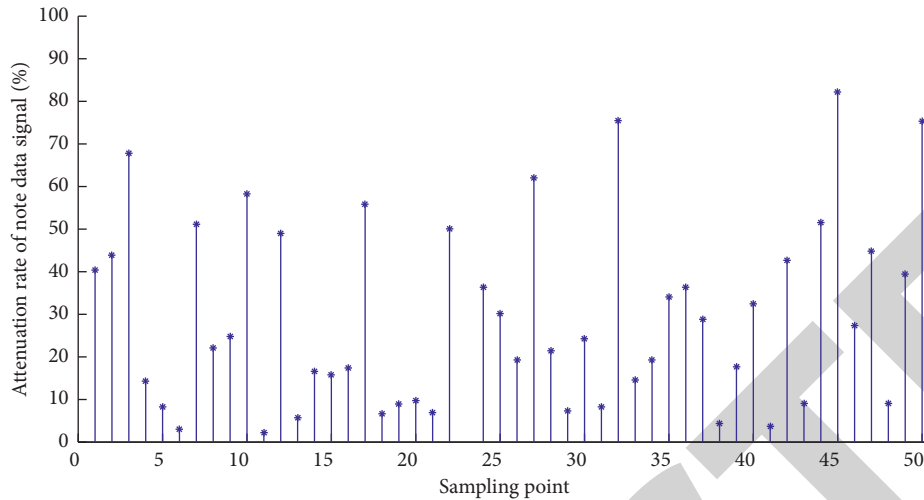
FIGURE 6: Match stick graph of attenuation rate of note data signal based on measurement learning.
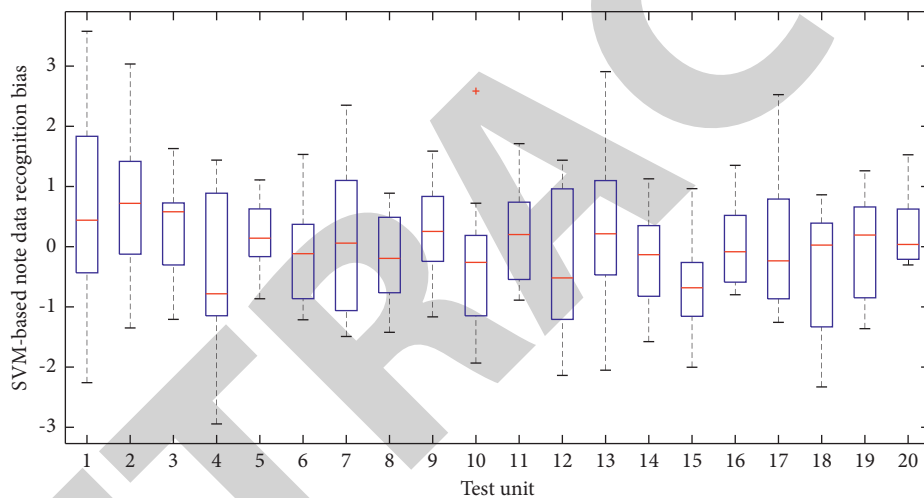
FIGURE 7: Box diagram of note data recognition deviation based on SVM.

From the knowledge of music theory, we also know that the different combinations of fundamental frequency and overtones are precisely the characteristics of the sound of an instrument, and it is the mark that distinguishes it from other instruments. Among them, the delay time and embedding dimension of the original signal are determined as 5 and 9, the delay time and embedding dimension of the first subsignal are determined as 2 and 7, and the delay time and embedding dimension of the second subsignal are determined as 4 and 9. The nonlinear characteristics of the training signal are used for the construction of the SVM classifier, and the nonlinear characteristics of the test signal are used for the recognition experiment. According to the recognition results, after adding empirical mode decomposition and SVM preprocessing in the recognition process, the recognition accuracy rate is better than the method of directly using the phase space of the note signal to reconstruct the matrix, and the average accuracy rate is improved.

3.3. *Example Results and Analysis.* The algorithms in this paper are all implemented with MATLAB software, and the computer environment used is a 32-bit Windows 7 operating system, with a built-in Intel I5-4200M processor and 4 GB of memory. In the experiment, in order to better evaluate the classification performance and generalization ability of the ML-SVM algorithm proposed in this paper, the ML-SVM algorithm is combined with logistic regression and linear kernel function in two musical score data sets of 9 and 4 difficulty. Each experiment was repeated 5 times independently, and a fivefold cross-validation was used, and the average accuracy rate was taken as the final recognition accuracy rate. At the same time, in order to evaluate the performance of the recognition algorithm more comprehensively, the 90% confidence interval of each algorithm result is also given.

Figure 8 shows the histogram comparison of the SNR (signal-to-noise ratio) of note data with the same algorithm. We use the B-spline frame twice to approximate the signal and the Gaussian frame to approximate the acoustic guitar d
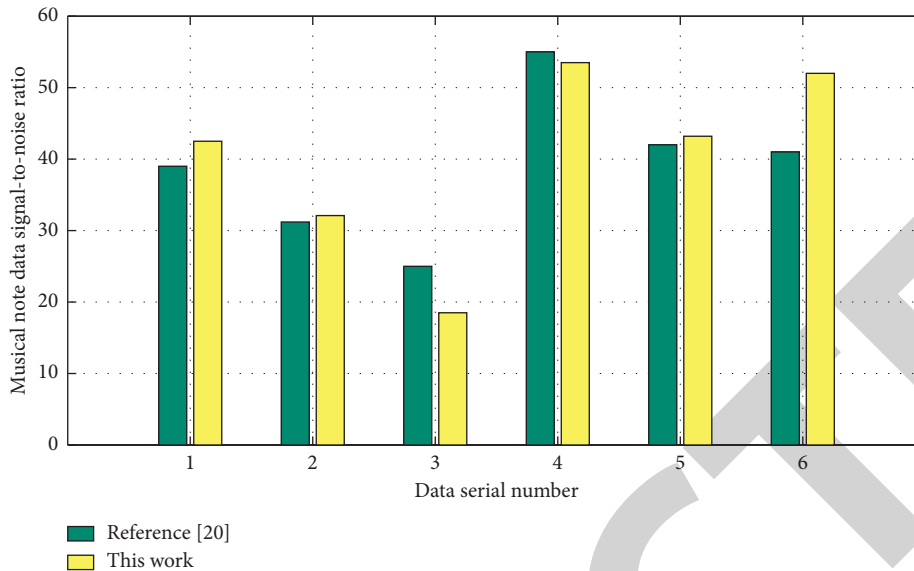
FIGURE 8: The comparison chart of SNR of note data of different algorithms.

1, and the error is measured. The characteristics of these three examples are clearly contrasted and can truly reflect the practical effects of the framework under various signal compositions. We see that, for signals with clear and clear frequencies, the result of frame approximation is very satisfactory.

In the actual playback of this signal, the error is very small, and human ears can hardly distinguish the difference. It can be seen that the error is larger at the beginning of the note and decreases with time. This is because plucked instruments are affected by factors such as plucking strength and angle, which will cause other strings at the beginning of the note. Resonance and the noise caused by flicking will therefore weaken over time.

Figure 9 shows the amplitude response curves of different note signal frequencies. In addition, when setting the music frequency, we only kept the 10 frequencies of the piano, with a maximum value of 1000 Hz. In actual recording, the effector will add harmony and simulation effects. These effects contain a lot of high-frequency signals, but these frequencies are discarded as noise. When the signal is actually played, the note restoration is highly recognizable, and the pitch and timbre information of the original signal is preserved, but it does not sound as bright and crisp as the original signal. This is precisely the result of discarding the high-frequency signal. We see that the response amplitude value is much lower than the previous two.

Therefore, if we know the frequency composition of a certain musical instrument playing each note, then for the calculated row vector in A, the combination of its nonzero elements in turn corresponds to a note. This is caused by the upgrade of signal complexity. Due to the influence of the sound of the instrument itself, the superimposition of notes is serious, and it is also affected by the recording environment and the structure of the piano itself. Environmental noise and high-frequency overtones are more and more complicated than the previous two examples. In actual playback, the approach signal retains the main frequency composition of the original signal, the note recognition is high, the energy feedback is obvious, and the beginning and end of the note can be clearly heard. However, part of the timbre information is lost, and the recognition of the instrument is reduced. This is related to the piano's own frequency response range. The threshold of the main frequency selection can be lowered to improve the approximation effect.

In order to improve the generalization ability, we perform multiparameter approximation on the training audio and obtain different frequency combinations of the same note by adjusting the Fourier transform threshold and the frame translation distance. Numerical experiments show that SVM can recognize single-tone melody very well, and every moment is correctly classified into the frequency that appears in the audio, and no nonaudio frequency is recognized, so the recognition rate of the note is 100%. However, due to the attenuation nature of the audio signal and the setting of the playing time, in the intersecting zone the previous note is about to stop and the next note has been played.

Figure 10 shows the line comparison chart of the recognition accuracy of musical note signal algorithms. It can be seen that, in the two data sets, among several algorithms, the final recognition accuracy of the algorithm proposed in this paper is higher than that in [19], in [20], and in [21], especially in the data set. The proposed algorithm obtains a recognition accuracy rate of 90.67%, which is greatly improved compared to the 85.63% literature algorithm. And a narrower confidence interval indicates that the result is more stable and significant. Compared with theoretical expectations, when the sampling rate is 1000, the duration error is between 0.5% and 1.6%. If the sampling rate is reduced, not only can the classification accuracy be improved, but also the real-time processing requirements for music within 200 tempos can be met.
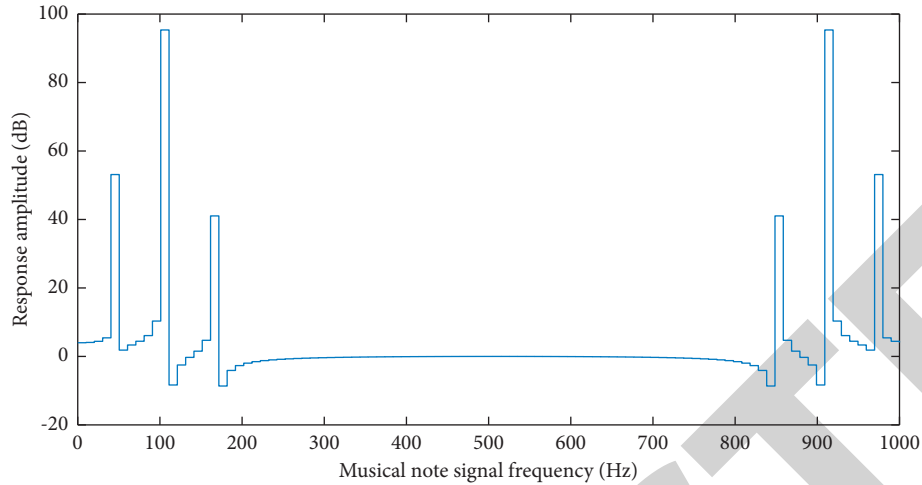
FIGURE 9: Amplitude response curves of different note signal frequencies.
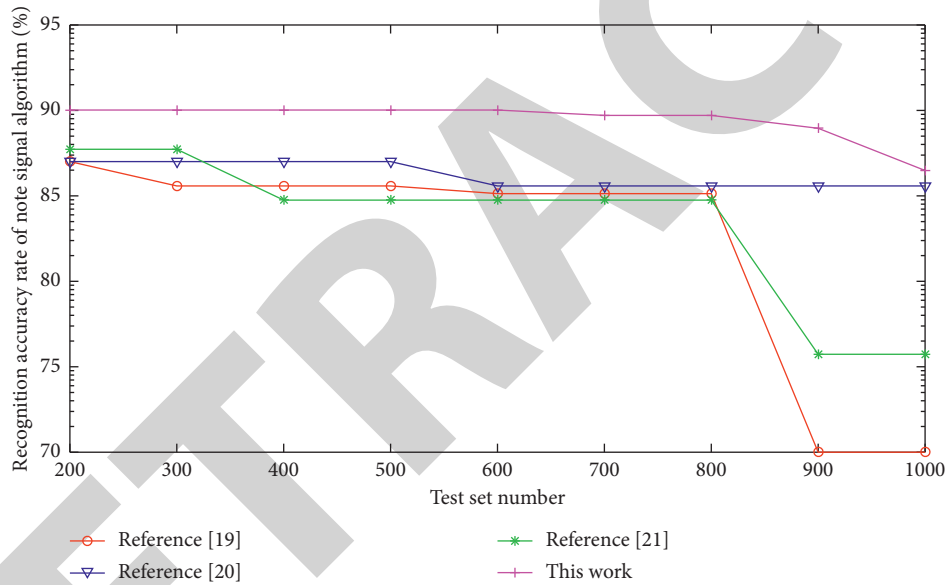
FIGURE 10: The comparison chart of the broken line of the recognition accuracy of the musical note signal algorithm.

4. Conclusion

The piano music note recognition system designed in this paper includes four main steps: audio input, spectrum preprocessing, feature calculation, and measurement learning support vector machine classification. Because the method of using fixed parameter values for different types of notes to extract the characteristic parameters will cause the loss of some nonlinear information of the notes and ultimately affect the recognition accuracy, this paper applies the method of adaptive signal decomposition to the signal preprocessing link. At the same time, in order to improve the recognition efficiency and reduce the time required for recognition, the paper also uses the SVM method to extract the principal components of the nonlinear features for the construction of the classifier and specific note recognition experiments. Then we propose a new method of

monophonic melody recognition based on frame approximation theory by using the accurate characteristics of the Fourier transform in the frequency domain, the frequency of energy concentration is detected in a large window, and then the idea of least squares is used to approximate the original signal, and on the basis of preserving the time-frequency characteristics of the original signal to the greatest extent, the audio is standardized and organized. Combining support vector machines based on statistical methods to recognize notes, it has the characteristics of supporting multiple frequencies and multiple notes, fast recognition, expandable models, and strong generalization capabilities. Among them, the spectrum preprocessing mainly uses the nuclear norm convex optimization technology to reconstruct the harmonic components in the original audio and adds a norm penalty term to enhance the robustness of the calculated spectrum against sparse noise; then the scale mapping is

Hindawi

Research Article

Multilayer Feature Extraction of AGCN on Surface Defect Detection of Steel Plates

Chi Zhang [ID],[1] Jian Cui [ID],[1] and Wei Liu [ID][2]

[1]*Department of Industrial Engineering and Management, Peking University, Beijing 100871, China*
[2]*China Mobile Research Institute, Beijing 100053, China*

Correspondence should be addressed to Chi Zhang; z.chi@pku.edu.cn

The development of industry is inseparable from the support of steel materials, and the modern industry has increasingly high requirements for the quality of steel plates. But the process of steel plate production produces many types of defects, such as roll marks, scratches, and scars. These defects will directly affect the quality and performance of the steel plate, so it is necessary to effectively detect them. Steel plate surface defects are characterized by their types, shape, and size: the same defect can have different morphologies, and similarities can exist between different defects. In this paper, industrial steel plate surface defect samples are analyzed, and a sample set is established by screening the collected defect images. Then, annotation and classification are performed. A multilayer feature extraction framework is developed in experiments to train a neural network on the sample set of defects. To address the problems of low automation, slow detection speed, and low accuracy of the traditional defect detection methods, the attention graph convolution network (AGCN) is investigated in this paper. Firstly, faster R-CNN is used as the basic network model for defect detection, and the visual features are jointly refined by combining attention mechanism and graph convolution neural network. The latter network enriches the contextual information in the visual features of steel plates and explores the semantic association between vision and defect types for different kinds of defects using the attention mechanism to achieve intelligent detection of defects, thus enabling our method to meet the practical needs of steel plate production.

1. Introduction

Steel is widely used in daily life and industrial production in a very large number of application areas and is the basic material for many products. According to statistics, in people's daily use of metal, steel accounts for up to 90% and is involved in most common products (home appliances, cell phones, etc.). The use of steel is essential in many industries, such as petrochemical, vehicle engineering, aerospace, military defense, and other fields, and its excellent performances have an invaluable role [1]. China's steel production in 2019 was as high as 996 million tons, accounting for 53.31% of the world's total production, far ahead of any other country [2].

Steel is an important and indispensable material in the modern construction of the country, and it is used in various production fields, especially in automotive, construction, and bridges. The production of steel plate is an extremely long and complicated process, from the raw stone to the final plate. It undergoes several processing steps, so the finished steel plate will inevitably have defects, most of them in the form of iron oxide, holes, cracks, scrapes, scratches, etc., on its surface. These defects directly affect the compressibility, toughness, corrosion resistance, and plasticity of the steel plates, rendering the manufactured products unable to meet customer requirements and resulting in severe economic losses for manufacturers. Without a set of effective testing methods, unqualified steel plate products put into use can even endanger people's lives and safety. Although steel production is high and export volumes are large, China still suffers from a slight shortage of automation compared to some developed countries [3]. Surface quality is a very important performance indicator of steel products, but it has not received the attention it deserves. According to statistics,

because of surface quality problems, large steel industrial groups lose an average of about 6 million U.S. dollars a year because surface defects lead to the return of the products. According to the market response, the vast majority of companies are challenged by the presence of surface defects, causing huge economic losses and hindering the transformation and upgrading of enterprises. Hence, controlling steel surface defects at the source is a very effective measure. Therefore, steel plate manufacturers need to carry out effective quality inspection of their products, to screen the steel plates that do not meet the technical specifications and improve the yield rate of the delivered products. On the other hand, analyzing the causes of surface defects will provide a reliable basis for improving the steel plate production technology. However, it is difficult to detect surface defects on steel plates. Production is made in very harsh environments, so it is difficult to install and protect inspection equipment; furthermore, moisture and impurities increase the difficulty of inspection [4].

Many researchers have studied defect detection and proposed several effective methods, mostly based on manual visual inspection [5], magnetic particle detection, penetration detection, eddy current detection, ultrasonic detection, traditional machine vision detection and identification, and deep learning detection and identification methods. But the recent developments in artificial intelligence theory and technology [6], the emergence of high-speed, high-precision CCD and CMOS industrial cameras, and the tremendous increase of CPU and GPU computing power and distributed computing provide the theoretical basis and hardware conditions for high-speed high-precision detection of steel surface defects based on computer vision. Deep learning benefits from the recent progress in computing power and automation technologies. It is one of the most representative fields of artificial intelligence, with excellent performance in image classification, target detection, segmentation, and target tracking. Deep learning stands out in the field of related technologies with powerful memory capability, nonlinear mapping capability, self-learning capability, and robustness [7]. Training a deep learning network with a large amount of data enables the low-level network to automatically learn the detailed features in the data and the high-level network to automatically learn the abstract features. So far, deep learning techniques have been widely used in agriculture, medicine, automotive, and aviation. One of its important applications in industrial production is product quality estimation, where it efficiently overcomes the shortcomings of traditional defect detection methods, more sensitive to human and external environment interference. Deep learning-based methods can detect products defect more quickly and accurately.

The surface quality of a steel strip is an important indicator to evaluate the grade of steel. Surface defects not only affect its appearance, which is not conducive to sales and exports but also affect its mechanical properties and quality, decreasing its stiffness and strength and reducing its corrosion resistance. Defects may also be the cause of serious safety accidents. Analysis of the steel surface defects shows that there can be various kinds of defects on the steel surface,

such as cracks, scratches, patches, inclusions, pitting, and bonds in the oxide skin, etc. In this paper, we propose a novel steel surface defect detection method (called AGCN) based on the attention mechanism and graph convolution neural network. The main contributions include the following: (1) the use of faster R-CNN [8] as the basic network model for surface defect detection and the combination of attention mechanism and graph convolution neural network; (2) exploring the contextual information in the visual features of steel plates and enhancing the semantic association between visual features and defect types using the attention mechanism and graph convolution neural network. Extended experiments are conducted on the steel surface dataset, and the advanced performance and effectiveness of the proposed method are demonstrated by method comparison and ablation analysis.

2. Related Work

2.1. Steel Plate Surface Defect Detection. The detection of steel surface defects began to develop in the 1920s and can be roughly divided into three families of methods: manual inspection, traditional photoelectric detection, and machine vision detection.

Manual inspection method, also known as manual visual method, was the first to be used. At that time, with the backward production technology, slow production speed, and low demand for products, the speed and quality of product testing were not very demanding. However, with the improvement in the level of production, the increase in demand and the disappearance of the demographic dividend, the shortcomings related to low inspection efficiency, high labor intensity, and nonuniform inspection standards gradually appeared and the inspection methods were no longer suitable for the requirements of speed and accuracy and were gradually abandoned [9].

Traditional photoelectric detection with optical sensors has been gradually applied in many enterprises, which improved the detection speed, accuracy, and efficiency [10]. Other methods based on eddy current detection and Faraday's electromagnetic induction principles have been developed. Common practice is to apply an alternating current to the strip surface, generating an alternating magnetic field affecting the detection coil. Measuring the induction current and impedance changes enables to determine the presence of defects. The main disadvantages are (1) waste of resources and (2) the method is not suitable for the detection of small defects. Leakage detection methods are based on the principle that when the steel plate is magnetized by a strong magnetic field, the change in the cross-sectional area of the steel plate due to the defect affects the magnetic permeability and reluctance. A part of the magnetic field bypasses the defective part through the surrounding air, causing the deformation of the magnetic field, from a straight line to a curve. The corresponding sensor converts the change of magnetic field into an electrical signal, and the size of the defect on the surface of the steel plate can be estimated from the electrical signal size. Leakage detection technology is a very simple, highly reliable, and fast detection technology,

subject to small environmental interference factors. But at the same time, there are simple leakage signal characteristics, detection of defects, and a limited variety of disadvantages [11]. Infrared detection methods use the principle that any object continuously emits infrared waves depending on its temperature, according to Planck's law of radiation, Wien's displacement law, Stefan–Boltzmann's law, etc. Infrared detection is more functional and can be used for defect detection, stress and fatigue analysis, simulation, image processing and fault diagnosis, and some other functions. The detection device is made of three main parts: the excitation part (signal generator and excitation source), the infrared camera, and the PC terminal. The working process is roughly to apply the corresponding excitation source to the steel material to obtain its thermal phase diagram. Since the temperature of the defective and nondefective parts of the steel is not the same, this will form areas with different temperature levels on the steel surface and emit different infrared waves. However, this method can only detect a few defect types, so it cannot be widely used. Laser scanning detection uses a laser as the emitting light source. The laser beam is reflected on the surface of the strip towards a rotating reflector and finally through the optical system equipped with an electric multiplier tube which converts it into an electrical signal through the converter, so that it can ensure the detection of defects at a certain speed. Using upper and lower lasers, this method enables to scan simultaneously both sides of the strip and obtain two-sided data of the inspected material. The final image is then processed and analyzed by a computer. However, laser detection technology requires high environmental protection, allows only slow detection speed, continuous motion of the laser life and reliability is reduced, and the purchase and maintenance costs are high. The limitations are large, and it is difficult to make further breakthroughs. So many companies began to look for low cost, high detection efficiency, and easy maintenance detection equipment.

With the breakthrough of CCD (charge-coupled device) camera and related hardware, while computer technology gradually developed and gradually appeared in industries, online machine vision-based steel plate inspection became popular. Figure 1 shows the working process and the role of each part: the steel plate to be inspected moves below the light source which provides additional light to remove the influence of background light and provide clear images collected by the camera which transmits them to the computer [12]. The computer analyzes and processes the images in real time according to the existing model and framework to assess the grade or defect situation of the steel plate surface. The use of CCD image sensors and pattern recognition technology has greatly improved the efficiency of steel surface defect detection, while various industries have started to detect surface defects with the help of similar devices.

Deep learning-based machine vision algorithms are the core of machine vision surface inspection technology, which is the key technology of the whole inspection system and one of the most challenging problems in the whole machine vision field. For steel surface defect detection, it is a popular research direction today to study algorithm models that can be executed with high accuracy, quickly extract the image features, and accurately identify the defect's category and location in real-time using massive image data. Deep learning is an end-to-end feature extracting algorithm, in which the model is similar to a black box. The process involves entering an image that contains a defect into the black box of the deep learning model which provides the category and location of the defect. Compared with traditional manual feature extraction, deep learning-based feature extraction enables more complete and accurate understanding of sample defects and features, thus achieving precise identification. Foreign research in this field began relatively early, for rail surface defect detection. The study and experiments of deep learning network models have concluded that different regularization methods have a certain impact on the recognition rate of defects. Sun et al. [13] designed a device for the identification of casting defects based on the mask R-CNN target detection model. Domestically, Cai and Wei [14] improved the YOLO target detection model with an accuracy of 97.55% in steel surface defect detection. He et al. [15] gave a multilayer feature fusion network structure using region proposal net (RPN) to generate regions of interest (RoI) on feature maps and obtained up to 82.3% of mAP (mean average precision) value on the dataset of defect detection at Tohoku University. To address the problem of insufficient dataset, Cui et al. [16] in foreign countries enhanced and expanded the dataset by cropping the original image, applying horizontal flipping, mirror flipping, transparency, and other processes. Liu et al. [17] used GAN (generative adversarial networks) network models to generate new defective dataset sample by merging original defect samples and defect-free samples and expand the dataset and achieve the purpose of sample migration. From the above analysis, with the increase of steel production, the defect detection of steel surface has stepped into the era of intelligent detection.

2.2. Multilayer Feature Extraction.

Convolutional neural networks (CNN) have developed rapidly and caught everyone's attention with their powerful modeling capabilities. Compared with traditional methods, the introduction of CNN has brought great improvements to areas such as image processing and natural language processing, for automatic translation, image, and speech recognition. However, traditional CNNs can only process Euclidean space data (e.g., images, text, and speech), which are translation invariant in these domains [18]. Translational invariance allows us to define convolution networks by defining globally shared convolution kernels in the input data space. Taking image data as an example, a picture can be represented as a set of regularly spaced pixels in the Euclidean space, and translation invariance means that a local structure of the same size can be obtained around any pixel [19]. Based on this, CNNs model local connectivity by learning convolution kernels shared at each pixel, which in turn create meaningful hidden layer representations for pictures. Although traditional CNNs bring enhancements in text and
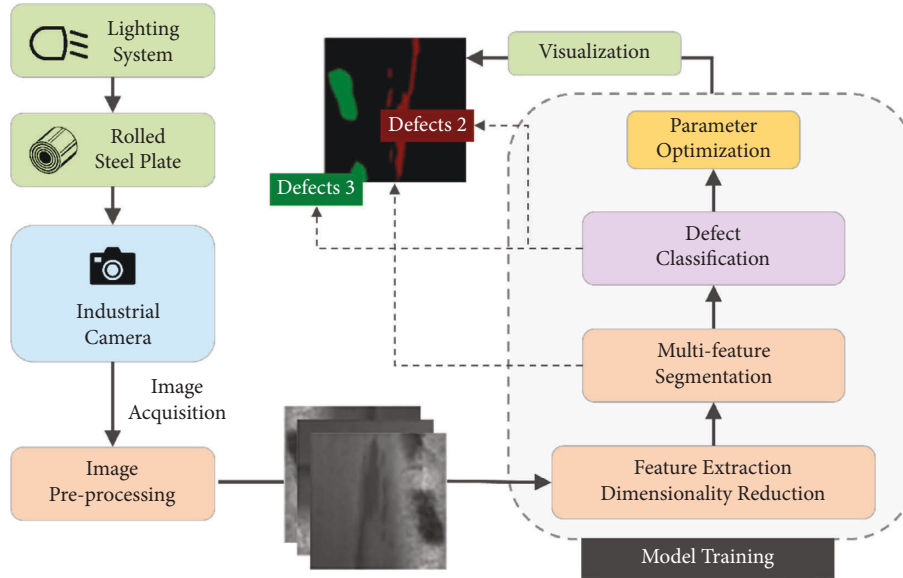
FIGURE 1: Process of computer vision diagnosis of surface defects on steel plates.

image domains, they can only handle Euclidean space data. Meanwhile, non-Euclidean spatial data—graph data—are gradually gaining attention due to its ubiquity [20].

In defect segmentation, the model needs to extract sufficient and effective semantic information to describe the difference between foreground (refer to the defects) and background (refer to the noise). Mahendran and Vedaldi [21] considered that convolution network feature maps in different layers contain the characteristics of different context information abundance. As shown in Figure 2, low-level feature maps with high resolution have clearer edges more detail information, which can be used to describe specific texture feature, but it contains less context information. On the contrary, the context information of high-level features is more abstract, and the semantic information is more separable after multilayer convolution extraction, but the texture details cannot be extracted due to low resolution. For classification problems, most methods mainly focus on high-level features, resulting in poor defect segmentation results in complex backgrounds. Inspired by the multilayer feature fusion method [22], this paper introduces the boundary refinement module to retain the low-level texture information.

Besides, the segmentation models need to be non-deformable for various variations such as defect's shape, scale, and texture. Most CNN-based methods try to expand the receptive field to cover the entire defect for global perception. In the DeepLab model [23], the receptive field is extended in the last convolutional layer to enhance the recognition of feature changes, but this will lead to grid artifacts [24]. Zhao et al. [25] use pyramid models with different pooling cores to amplify local features to overcome intraclass differences. However, excessive pooling in feature fusion makes the model unable to capture a wider range of global information, resulting in missing parts when marking defect masks [26]. To solve the above problems, we propose a multilayer feature fusion method, which uses multiscale

convolution (receptive fields of different sizes) to weight the feature maps of all convolutional layers to obtain the context information. On the premise of fully exploiting defect features, grid artifacts and excessive pooling are avoided.

2.3. Graph Convolution Neural Network.

Graph data can naturally represent real-life data structures, such as traffic networks, World Wide Web, and social networks. Unlike image and text data, the local structure of each node in graph data varies, which makes translation invariance no longer satisfied. The lack of translation invariance poses a challenge to define CNN on graph data. In recent years, due to the prevalence of graph data, researchers have started to focus on how to construct deep learning models on graphs. With the ability of CNNs to model local structures and the prevalence of node dependencies on graphs, GCN (graph convolution neural) networks have become one of the most active and important research fields. Recently, several articles have been published to explore deep learning on graphs, but there is still a gap in the in-depth discussion and summary of the modeling methods and applications of the most important branch, graph CNNs. In this paper, we summarize the development of GCNs and their future trends [27].

The challenges faced in the construction of GCNs are mainly related to the following aspects: (1) graph data are non-Euclidean spatial data and do not satisfy translation invariance, i.e., each node has a different local structure. The basic operators in traditional CNNs (convolution and pooling) rely on the translation invariance of the data. At this point, it becomes a challenging task to define convolution and pooling operators on graph data. (2) A variety of real-life applications can be naturally represented by graphs, which give them diverse properties, such as directed connections of users in social networks, heterogeneous connections of authors and citations in citation

Low-level
Features

High-level
Features
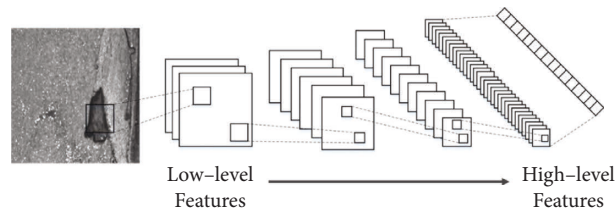
Figure 2: Feature extraction from low level to high level.

networks, and positive and negative tendency band symbolic connections in political relationship networks. The various graph characteristics bring more information to the construction of GCNs, but the modeling of multiple characteristics also requires a more complex and detailed design of GCNs, which brings new challenges. (3) The scale of graph data is very large: in the era of big data, graphs in practical applications may be extremely large, with millions or even tens of millions of nodes, such as user commodity networks in recommendation systems and user networks in social networks. It is very challenging to build GCNs on large-scale graphs in an acceptable range of time and space [28].

In addition, researchers borrowed knowledge from graph theory, such as using eigenvalues and eigenvectors of Laplacian matrices for community analysis or population clustering. With the rise of deep learning, researchers started to consider introducing deep learning models into graph data, and the representative research work is called network embedding, i.e., learning fixed-length expressions for each node by constraining the proximity of nodes. This led to new methods such as Deep Walk, LINE, and node2vec. During this period, when solving specific application problems, researchers usually modeled them as two-stage problems [29]: taking node classification as an example, the first stage learns uniform-length expressions for each node, whereas the second stage uses node expressions as inputs to train classification models. In recent years, researchers have gradually shifted their focus from modeling graph data to how to migrate deep learning models to graphs for end-to-end modeling, and GCNs are one of the most active fields. In modeling graph convolution neural networks, researchers focus on how to build convolution operators on graphs. Zhang et al. [30] proposed the first graph convolution neural network in 2013, where they defined graph convolution in the spectral space based on graph theory using the convolution theorem. This branch was later developed as the spectral approach in the field of graph convolution. The initial spectral methods had the disadvantage of high spatio-temporal complexity, and Cheb-Net and GCN parametrized the convolution kernel in the spectral domain to greatly reduce the spatio-temporal complexity [31]. These two methods, although categorized as spectral methods, have started to define the weight matrix of nodes from a spatial perspective. Inspired by these two methods, spatial methods were applied and began to consider modeling the weights between nodes in the node domain with attention mechanisms, serialization models, etc. The graph convolution neural networks of this period did not take too much

account of the characteristics of graphs in the process of constructing convolution operators. With the gradual improvement of convolution operators, people began to consider various features of graphs, starting with a focus on how to model higher-order information on graphs, and fine-grained designs for graphs with features on edges, heterogeneous graphs, etc. In addition, the question of how to train more efficient GCNs has also received much attention. Researchers have started to train deeper GCNs to enhance generalization. In addition, the scalability of the models to large-scale graphs and the training speed are very focused research directions in GCN. The pooling operator, as the main component of CNNs, enables to expand the perceptual field and reduce the number of parameters. Recently, some research has also started to focus on the construction of on-graph pooling operators [32]. The on-graph pooling operator is mainly used in graph classification problems with the aim of learning the hierarchical structure of the graph. The broad application scenarios targeted by graph data modeling makes the tasks handled by graph data modeling diverse. We divide the downstream tasks into node-level tasks and graph-level tasks. Node-level tasks include node classification and link prediction, such as article classification in citation networks and inference of user preferences for products in recommendation systems. Graph-level tasks include graph generation and graph classification, such as drug network generation and protein classification in protein networks [33].

3. Method

3.1. Steel Plate Defect Analysis. Since the original size of each defect can vary greatly, we will apply scaling to achieve a uniform size. Figure 3 shows five common steel plate surface defects: their basic characteristics are as follows. (1) White iron scale: mainly in the form of strips of varying length, the color is generally white, mostly in patches of aggregated distribution, and the size of the defect varies. (2) Roll marks include three main types of features: defects for the lighter gray-white distribution of scattered microarcs, with low contrast to the background; defects for the dark gray arc-shaped microfolds; and a small number of defects for the continuous gray-black periodic straight band. (3) Scratches are generally gray-black, mainly in the width of the continuous periodic band. When the background color is dark, the defect looks like it and the contrast is low. (4) Scarring: mainly in the form of black dots or surfaces of different sizes, usually aggregated in patches, part of the distribution is more scattered. The background brightness may vary: as the

(a)   (b)   (c)   (d)   (e)
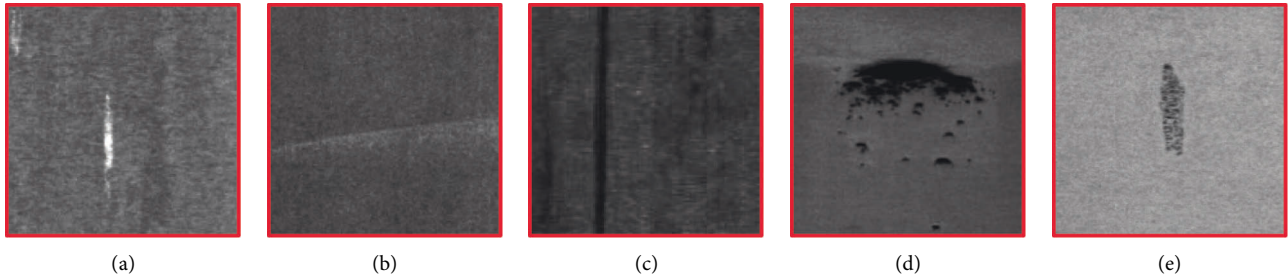
FIGURE 3: Types of defects in the collected data set. (a) White iron scale. (b) Roll printing. (c) Scratch. (d) Scarring. (e) Embroidery skin.

background gets darker, the contrast of features gradually decreases. (5) Rusty skin/embroidery skin. Mainly a certain width of short gray-black bands, features are more obvious, usually a single distribution or a very small number of clustered distributions. The above five kinds of defects are the ones that are the most studied.

The analysis of a large number of defect images shows that there are similarities between different types of defects (such as interrupted parts of rusty skin and scars, shallower, smaller scars, and white iron scales), and the same type of defects have a variety of forms and sizes (roll printing defects have three different forms and sizes). Traditional vision inspection methods have difficulty solving these problems, while GNN-based inspection algorithms can effectively detect these complex forms of defects.

3.2. Model Architecture. The proposed model is shown in Figure 4 and is made of three parts: (1) multilayer feature extraction network as backbone for steel plate defect detection, to extract visual features and spatial information of salient regions. (2) Graph CNN: to enrich the contextual information of visual features. (3) Attention mechanism: to explore the semantic association between visual features and fault categories.

3.3. Multilayer Feature Extraction. The multilayer feature extraction module is divided into 4 parts, which is feature extraction network, RPN (region proposal network) module, RoI pooling module, and R-CNN module [34]. It mainly generates candidate regions and performs preliminary classification and localization through RPN [35]. Then, it pools the acquired candidate regions and finally classifies and again improves the positions of the pooled defect features.

In this paper, firstly, we establish the global context attention mechanism into the adjacent resolution feature map. Secondly, the global context information is extracted from the low-resolution feature map. Thirdly, the high-resolution information is weighted to refine the spatial position of the category pixels, which can ensure the high-level features are not weakened and achieve a more accurate classification result without increasing the amount of calculation. Consider RPN network can map the generated region to the feature map generated by the convolution network through "anchors," realizing the connection

between the two and further improving the detection speed and accuracy [36]. To learn whether a defect is present in the input image, anchors (rectangular boxes with a certain size and aspect ratio) are placed on the image for each location on the output feature map from the RPN network. Then, the anchor is matched with the real defect, and the classification and fine-tuning of the defect location is performed. Figure 5 shows the computational flowchart of the multilayer extraction process, dividing the detection process in two steps and providing preliminary localization and classification (proposal). The more accurate the proposal is, the smaller the error of the later redetection. The RoI is obtained by screening a large number of proposals generated by the predicted anchor during training, and the proposal is directly used as the RoI during testing. RPN module mainly consists of five submodules:

(1) Anchor generation: RPN corresponds to nine anchors for each point on the generated feature map, and each anchor has three different area sizes and three different aspect ratios, corresponding to the original map covering possible defects.

(2) RPN convolution network: by employing a convolution network, each generated anchor is processed to obtain its prediction score and offset value.

(3) Calculate RPN loss: this part occurs only during the training process, matching the anchor with the labels to distinguish between positive and negative samples, obtaining the true values of classification and offset, and calculating the loss with the prediction score and offset values obtained in the previous step.

(4) Generate proposal: screen the anchor obtained by the RPN convolution network to get a better set of proposal for the subsequent network.

(5) Screening proposal to get RoI: screen the proposal obtained in the previous step to get the final RoI.

3.3.1. Feature Extraction Network. In order to obtain a better feature map if the image has a low contrast, a 13-layer convolution network similar to VGG16 is used to extract the defect features of the scaled image. The convolution layers use a small convolution kernel of size 3×3 with a small number of parameters, while the large number of layers gives it a better nonlinear capability to improve its learning
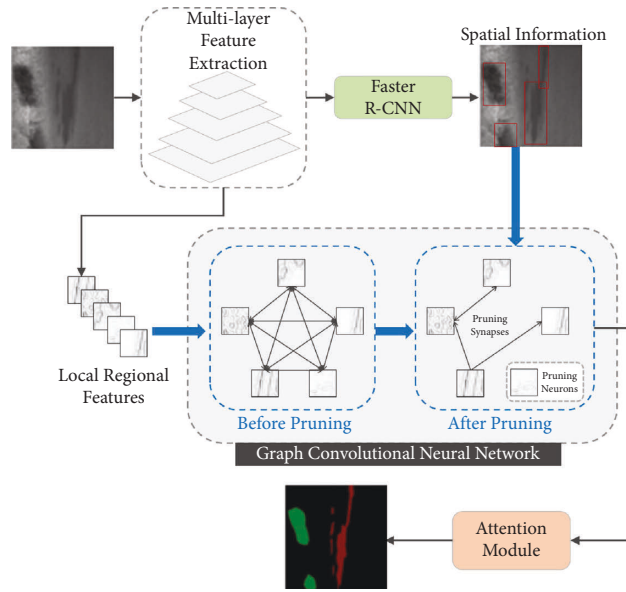
Figure 4: Model architecture.
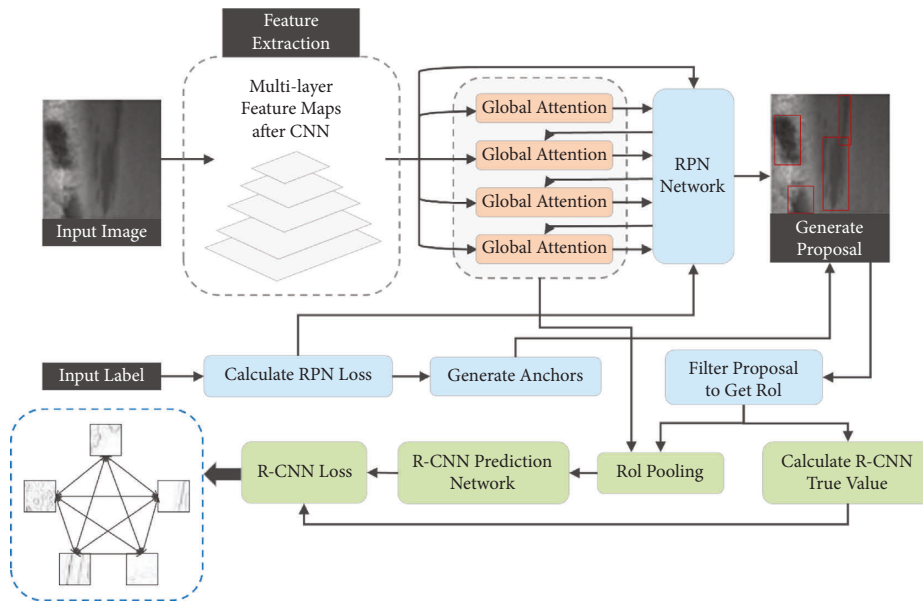
Figure 5: Schematic diagram of the multilayer feature extraction process.

capacity. This part includes five convolution modules, each outputting feature maps of 64, 128, 256, 512, and 512, respectively.

3.3.2. RPN Module.

Anchor generation is performed on the feature maps, the category score, and position offset value of each anchor are predicted, and the binary classification of defects (i.e., the presence or absence) and their preliminary location are performed according to the acquired proposals [37]. This module shares the convolution features of the whole map with the R-CNN detection network, saving time and providing high quality proposals to the R-CNN detection network, which improves the detection accuracy of the model. The number of output feature maps is 512, and 18 and 36 in the later classification and localization parts, respectively.

3.3.3. RoI Pooling Module.

Since the R-CNN module behind the faster R-CNN uses a fully connected layer, a uniform dimension is required before the defect feature subgraphs are input to the fully connected layer (the feature subgraph size in this paper is 7×7). Since the defect images of steel plates are of different sizes, their RoI corresponding feature map sizes are also different. RoI pooling is used to perform feature scale transformation to enable taking defect images of arbitrary sizes as input and output a fixed feature map size,

suitable for detecting the five different defects scales of this paper. The number of feature maps in this part is kept constant.

3.3.4. R-CNN Module.

The RoI feature subgraphs obtained from RoI pooling are mapped to the whole feature map and output to the fully connected network of the R-CNN module, which performs the defect class prediction and location regression. This part of the fully connected layer outputs 2048 feature maps. Faster R-CNN outputs the visual features of the i-th anchor point as $h = \{h_i\}_{i=0}^{m} \in R^{m \times 2048}$ and the spatial information as $S_i = [x_{i1}, y_{i1}, x_{i2}, y_{i2}]$, where m denotes the number of salient regions.

3.4. Graph Convolution Neural Network.

A first fully connected graph is constructed and further refined using the accurate contextual information between multiple salient regions, to obtain a spatial graph network. We use faster R-CNN to obtain the visual features and spatial information of saliency regions, as shown in Figure 6(a). Next, we introduce the construction of the fully connected graph by considering each object region in the image as a vertex, and by constructing a relationship graph, we obtain a fully connected indirect graph as shown in Figure 6(b), where each edge represents the relationship between two regions. The spatial information of the regions representing the position of the regions in the image is a four-dimensional spatial vector $S_i = [x_{i1}, y_{i1}, x_{i2}, y_{i2}]$, where (x_{i1}, y_{i1}) is the coordinate of the upper left corner of the bounding box and (x_{i2}, y_{i2}) is the coordinate of the lower right corner of the bounding box. The identification of the correlations between regions is done according to the following steps: (1) the visual features of the two regions are fed into the multilayer perceptron to obtain feature integration, and the corresponding elements of the two-feature embedding are multiplied to obtain a correlation score. (2) We determine whether there is a correlation between two regions based on the size of the overlap area. If two regions have a large overlapping area, it means there is a strong correlation between these two regions. If the two regions do not have any overlapping part, we consider that these two regions have weak correlation, which means there is no edge connecting these two nodes. In addition, we identified five different categories of region relationships, such as internal, overlay, and overlap. Based on the spatial relations, we removed some irrelevant region relations from the fully connected graph and obtained a relatively sparse graph as shown in Figure 6(c).

To enhance the contextual information in the visual features of each region, we use GCN to update the object representation. If the image contains m salient regions, considered as nodes, we use the $m \times m$ adjacency matrix A to represent the structure of the graph, with $A_{ij} = 1$ if there is an overlapping region between node i and node j; $A_{ij} = 0$ otherwise. Given the target node i and neighboring node $j \in N(i)$ in the image, where $N(i)$ is the set of nodes adjacent to node i, the visual feature representations of node i and node j are h_i and h_j, respectively. To obtain the correlation score s_{ij} between nodes i and j, by splicing the visual features of h_i and h_j, we first train a fully connected layer.

$$s_{ij} = w_a^T \sigma\left(W_a[h_i, h_j]\right), \quad (1)$$

where w_a and W_a are the learning parameters, σ is the nonlinear activation function, and $[h_i, h_j]$ denotes the concatenation operation. We apply the softmax function on the correlation score s_{ij} to obtain the weight α_{ij}, as shown in Figure 6(c).

$$\alpha_{ij} = \frac{\exp(sij)}{\sum_{j \in N(i)} \exp(sij)}. \quad (2)$$

For the graph convolution, the neighboring nodes of $h_j, j \in N(i)$ are first passed through a learned linear transformation W_b. These transformed representations are aggregated by the weights α_{ij}, and finally the updated node features h_{vi} are obtained by the activation function σ:

$$h_{vi} = \sigma\left(h_i + \sum_{j \in N(i)} A_{ij} \alpha i_j W_b h_j\right). \quad (3)$$

The output feature of the last layer node i in GCN is H_i, and the set of features of all nodes is H.

3.5. Attention Mechanism.

In order to enhance the higher-order semantic association between visual features and defect types, to refine and reduce the redundant information in visual features and highlight the key semantics, we designed the attention mechanism. First, for the regional visual representation H obtained by convolution of the graph neural network, each node's feature set is updated using the self-attention mechanism to obtain a new feature set $\tilde{H}$:

$$\tilde{H} = \mathrm{softmax}\left(\frac{HH^T}{\sqrt{d}}\right)H, \quad (4)$$

where H^T is the transpose of H and d is the dimension of H. To obtain the visual representation associated with the surface defect representation of the steel plate, we utilize the design learnable parameter matrices W_m. W_m is used as a guiding matrix to adjust the visual representation $\tilde{H}$. The similarity score between W_m and $\tilde{H}$ is calculated as follows:

$$r = \frac{W_m \tilde{H}^T}{\sqrt{d_m}}. \quad (5)$$

For the i-th region, a softmax function is used to normalize the score ri to obtain the defect diagnosis probability $\tilde{r}_i$.

$$\tilde{r} = [\tilde{r}_1, \tilde{r}_2 \ldots, \tilde{r}_i] = \frac{\exp(r_i)}{\sum_{j \in m} \exp(r_j)}. \quad (6)$$

3.6. Loss Function.

We deploy a multitask loss function that combines classification loss and edge localization regression loss for unified training and finally outputs the

(a)　　　　　　　　　　(b)　　　　　　　　　　(c)
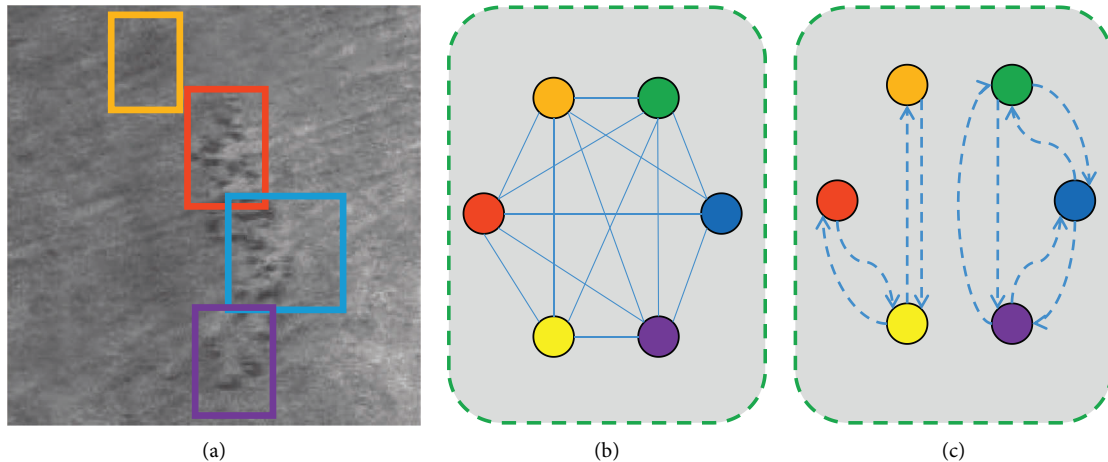
FIGURE 6: Graph generation.

corresponding classification and edge locations, which can improve the detection accuracy and is suitable for detecting small target defects:

$$L(\{\tilde{r}_i\}, \{S_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(\tilde{r}_i, r_i^*)$$
$$+ \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(S_i, S_i^*), \tag{7}$$

where i is the index of the anchor; $\tilde{r}_i$ is the probability of anchor being predicted as a target; r_i^* is the category label, 1 when a target exists and 0 otherwise; S_i is the predicted value of location regression; S_i^* is the label value of location regression; λ is the weight balance value; N_{cls} and N_{reg} are the normalized values of classification loss L_{cls} and regression loss L_{reg}, respectively.

4. Experimentation and Evaluation

4.1. Datasets. We set up a test bench in a steel production factory and used a line-scan camera to acquire online images of steel plates and obtained a total of 5,000 defect samples after screening, 1,000 for each defect type, of which 4,000 were used as training samples and 1,000 were used as test samples. In the datasets of this paper, each defect image contains at least one defect, and some images contain multiple defects of different scales to ensure that the trained detection model can adapt to complex detection [38]. In this paper, the defect samples are labeled with defects using labeling software with rectangular boxes, where the label name is the first letter of the Chinese pinyin capitalization of each defect, such as TL for white iron scale, GY for roll marks, GH for scratches, JB for scars, and XP for rusty/embroidery skin.

4.2. Experimental Setup and Evaluation Metrics. The experiments were completed using the conditions shown in Table 1. To further demonstrate the effectiveness of the proposed method, 6 models and the AGCN were trained for

TABLE 1: Experiments conditions.

CPU	Intel Xeon W-2135
Operating system	Ubuntu 18.04
RAM	32 GB
GPU	GeForce RTX 2080Ti
Video memory	8 GB
Python version	3.6.9
CUDA	10.0
CU DNN	7.4.1

100 epochs (rounds) on the datasets of this paper. The training methods are as follows: the initial learning rate of the model is set to 0.0001; the Adam optimizer is used; the learning rate decay is performed once every 5 epochs, and its decay rate is 0.1; the batch size is set to 8 (the batch size is the number of samples selected for one training epoch; it is limited by the GPU of the device, and chosen to obtain the best optimization and highest training speed). The loss value plot of the proposed model trained on the dataset of this paper with 100 epochs is shown in Figure 7, where the vertical coordinate is the loss value (loss) and the horizontal coordinate is the number of training sessions. The loss values converge quickly during the training process and finally sets to about 0.18.

4.3. Performance Comparison. In this paper, the average accuracy mIoU (mean intersection over union) and frame rate are used as the actual metrics for defect detection, and the final saved optimal network model is taken for testing. The average detection accuracies of the five defects on the test set are shown in Table 2.

In order to avoid overfitting, this paper adopts 5-fold cross-validation. Compared with the original method (faster R-CNN) and recent advanced methods, our method has the highest average mIoU of 0.8580. Among the existing five categories of defect recognition, our methods are the most prominent in the 4 categories of defects. Compared with the most accurate RefineNet model, we are
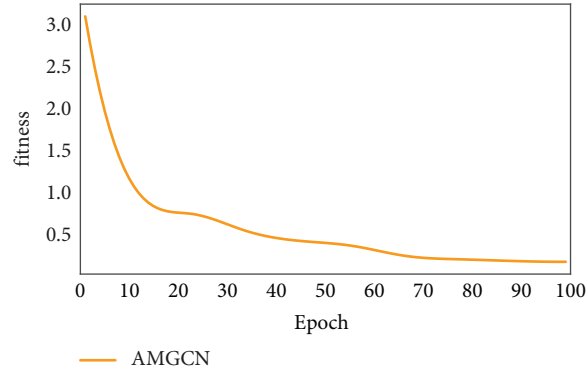
FIGURE 7: Training loss convergence plot.

TABLE 2: Performance comparison (mIoU).

Model	White iron scale	Roll printing	Scratch	Scarring	Embroidery skin	Average
Faster R-CNN [8]	0.7962	0.7204	0.8386	0.7602	0.8982	0.8027
SegNet [39]	0.8250	0.6835	0.8532	0.8622	0.8824	0.8213
PSPNet [40]	0.8032	0.7282	0.8419	0.8358	0.9047	0.8228
YOLOv4 [14]	0.8068	0.7025	0.8793	0.8524	0.8856	0.8253
DeepLab+ [23]	0.8271	0.7139	0.8748	0.8437	0.8754	0.8270
RefineNet [41]	0.8265	**0.7280**	0.8786	0.8613	0.8410	0.8271
AGCN (ours)	**0.8485**	0.7188	**0.9098**	**0.8762**	**0.9367**	0.8580

TABLE 3: Average detection speed.

Model	Testing time in seconds	Number of frames per second
Faster R-CNN	0.0621	16.11
SegNet	0.0464	21.54
PSPNet	0.0427	23.42
YOLOv4	0.0340	**29.39**
DeepLab+	0.0386	25.92
RefineNet	0.0534	18.73
AGCN (ours)	0.0364	**27.46**

only 0.0092 behind in the defect detection of roll printing. It is proved that this method is effective for steel plate defect detection.

On the other hand, the recognition accuracy of each category of defects is different. Since embroidery skin is generally reflective which have large contrast with the background noise, its features are easy to be captured by the algorithm, and its recognition accuracy is generally higher than that of other categories. For roll printing, it typically appears as a long strip. However, the length of such defects varies in different images. One picture probably has more than one roll printing defects, which leads to low mIoU.

From Tables 2 and 3, it can be seen that faster R-CNN obtains a better detection accuracy than YOLOv4 on the test set; the defect detection accuracy of AGCN is better than that of faster R-CNN. AGCN improves the average detection accuracy of embroidery skin defects to 97.67%, which is 7.5% better compared with faster R-CNN. Compared with faster R-CNN and YOLOv4 model, the detection time of AGCN model is increased, but the impact is not significant, and our model can meet the practical requirements. The trained

model is tested on the test set and found to have issues such as false detection and missed detection. Defect characteristics vary greatly between classes, and when several types of defects have similarities, the accuracy is low. If a single "White iron scale" defect is present in a small area of an image, after multiple convolutions, the model losses the defect characteristics and cannot detect it. For "Roll marks" defects, when the color is light, bright or white, this type of defect can be misdiagnosed as "White iron scale"; because when the color is very similar to the background, the characteristics are not obvious and detection can be missed. When several small target defects are detected in an image, such as rusty/embroidery skin defects, the AGCN model has significantly higher average detection accuracy (The segmentation map are shown inFigure8).

4.4. Ablation Experiments. To verify the role of the proposed attention module and the graph convolutional neural network, we did some ablation experiments. The number of convolutional layers of the graph convolutional neural
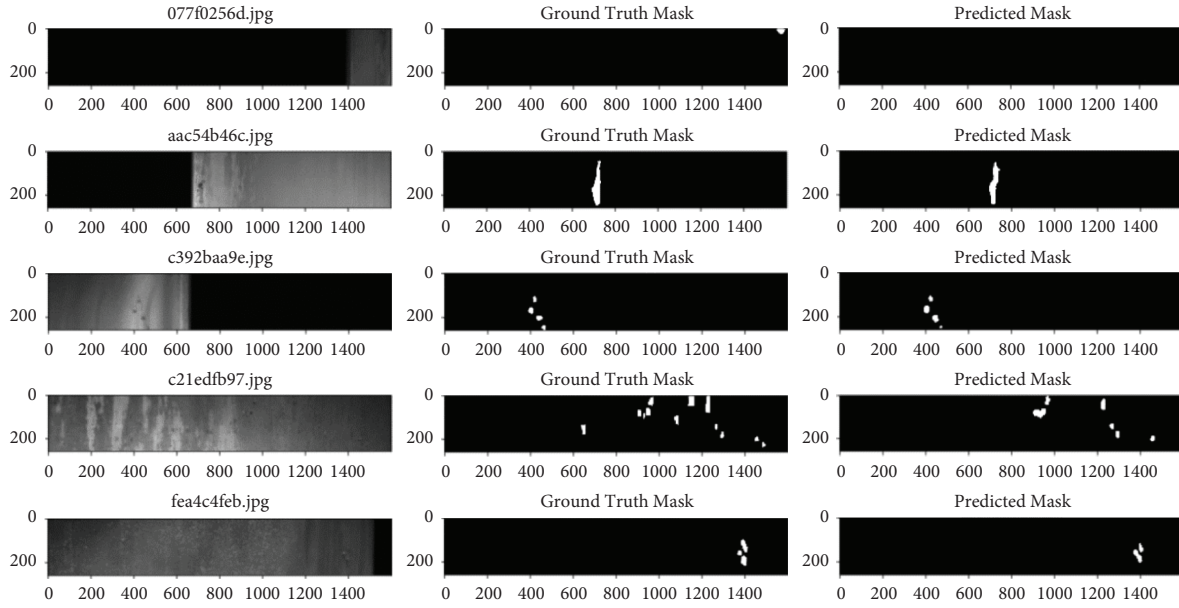
FIGURE 8: Partial experimental results—segmentation map.

TABLE 4: Ablation experiments.

Number of convolutional layers	White iron scale (%)	Roll printing (%)	Scratch (%)	Scarring (%)	Embroidery skin (%)
1	82.72	72.11	90.87	89.08	90.89
2	82.75	72.24	90.86	89.32	93.82
3	**84.85**	**72.88**	**90.98**	**89.62**	**97.67**
4	83.62	72.54	90.76	89.42	95.82

network is varied to verify the effect of the model. We set the number of convolution layers as 1, 2, 3, and 4. The results are shown in Table 4. The model achieves the best results when the number of convolution layers is 3, which indicates that there are deep semantic associations between different salient regions in the images that require multilevel graph convolution. The effectiveness of the proposed graph convolutional neural network is demonstrated.

5. Conclusion

This paper introduces a AGCN detection method based on the multilayer feature. We analyze the features of the most common steel plate defects and the characteristics of the faster R-CNN network-based model as the visual multifeature encoder and propose to refine the visual features by combining the attention mechanism and the graph convolution neural network approaches. The graph convolution neural network enriches the contextual information in the visual features of steel plates and further explores the semantic association between vision and defect types using the attention mechanism to achieve intelligent defect detection. This method meets the practical needs of defect detection in steel plate production. In the future, we plan to explore the dynamic neural network-based steel plate defect detection.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Chi Zhang was the initiator of the subject and completed article writing and first revision, put forward the model framework, and finished the experimental verification. Jian Cui supplemented the experiment, improved the attention mechanism, and completed the second revision according to the opinions of reviewers. Wei Liu added comparative experiments and integrated the concept of multilayer fusion into the model. At the same time, she embellished the whole text to make the article more compact.

Acknowledgments

References

[1] Y. He, K. Song, Q. Meng, and Y. Yan, "An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1493–1504, 2020.

[2] D. He, K. Xu, and P. Zhou, "Defect detection of hot rolled steels with a new object detection framework called classification priority network [J]," *Computers & Industrial Engineering*, vol. 128, pp. 290–297, 2019.

[3] H. Dong, K. Song, Y. He, J. Xu, Y. Yan, and Q. Meng, "PGA-net: pyramid feature fusion and global context attention network for automated surface defect detection," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7448–7458, 2020.

[4] Q. Luo, X. Fang, L. Liu, C. Yang, and Y. Sun, "Automated visual defect detection for flat steel surface: a survey," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 3, pp. 626–644, 2020.

[5] H. Di, X. Ke, Z. Peng, and Z. Dongdong, "Surface defect classification of steels with a new semi-supervised learning method," *Optics and Lasers in Engineering*, vol. 117, pp. 40–48, 2019.

[6] S. Tian and K. Xu, "An algorithm for surface defect identification of steel plates based on genetic algorithm and extreme learning machine," *Metals*, vol. 7, no. 8, p. 311, 2017.

[7] J. Li, Z. Su, J. Geng, and Y. Yin, "Real-time detection of steel strip surface defects based on improved YOLO detection network," *IFAC-PapersOnLine*, vol. 51, no. 21, pp. 76–81, 2018.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[9] K. Liu, H. Wang, H. Chen, E. Qu, Y. Tian, and H. Sun, "Steel surface defect detection using a new haar–weibull-variance model in unsupervised manner," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 10, pp. 2585–2596, 2017.

[10] Y. Jia, K. Liang, P. Wang, K. Ji, and P. Xu, "Enhancement method of magnetic flux leakage signals for rail track surface defect detection," *IET Science, Measurement & Technology*, vol. 14, no. 6, pp. 711–717, 2020.

[11] J. Yang, S. Li, Z. Wang, H. Dong, J. Wang, and S. Tang, "Using deep learning to detect defects in manufacturing: a comprehensive survey and current challenges," *Materials*, vol. 13, no. 24, p. 5755, 2020.

[12] M. M. Saari, N. A. Nadzri, A. M. Halil et al., "Design of eddy current testing probe for surface defect evaluation," *International Journal of Automotive and Mechanical Engineering*, vol. 16, no. 1, pp. 6357–6367, 2019.

[13] X. Sun, J. Gu, S. Tang, and J. Li, "Research progress of visual inspection technology of steel products—a review," *Applied Sciences*, vol. 8, no. 11, p. 2195, 2018.

[14] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution [J]," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[15] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "Lightgcn: simplifying and powering graph convolution network for recommendation [C]," in *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pp. 639–648, China, July 2020.

[16] Z. Cui, K. Henrickson, R. Ke, X. Dong, and Y. Wang, "High-order graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting [R]," in *Proceedings of the Transportation Research Board 98th Annual Meeting*, Washington DC, USA, January 2019.

[17] X. Liu, F. Gao, Q. Zhang, and H. Zhao, "Graph convolution for multimodal information extraction from visually rich documents [J]," 2019, https://arxiv.org/abs/1903.11279.

[18] Y. LeCun, B. Boser, J. S. Denker et al., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[19] J. Atwood and D. Towsley, "Recent advances in convolutional neural networks [J]," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[20] Y. Zhang, P. Qi, and C. D. Manning, "Graph convolution over pruned dependency trees improves relation extraction [J]," 2018, https://arxiv.org/abs/1809.10185.

[21] A. Mahendran and A. Vedaldi, "Understanding deep image representations by inverting them [C]," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5188–5196, Boston, MA, USA, June 2015.

[22] P. Zhang, W. Dong, H. Lu, H. Wang, and B. Yin, "Learning uncertain convolutional features for accurate saliency detection [C]," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 212–221, Venice, Italy, October 2017.

[23] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[24] P. Wang, P. Chen, Y. Yuan et al., "Understanding convolution for semantic segmentation [C]," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pp. 1451–1460, Lake Tahoe, NV, USA, March 2018.

[25] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network [C]," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2881–2890, Las Vegas, NV, USA, July 2016.

[26] R. Zhang, "Making convolutional networks shift-invariant again [C]," in *Proceedings of the International conference on machine learning*, pp. 7324–7334, Long Beach, CA, USA, June 2019.

[27] H. Lei, N. Akhtar, and A. Mian, "Spherical kernel for efficient graph convolution on 3d point clouds [J]," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3664–3680, 2021.

[28] Y. Liu, M. Zhang, X. Yin, Z. Huang, and L. Wang, "Debonding detection of reinforced concrete (RC) beam with near-surface mounted (NSM) pre-stressed carbon fiber reinforced polymer (CFRP) plates using embedded piezoceramic smart aggregates (SAs)," *Applied Sciences*, vol. 10, no. 1, p. 50, 2019.

[29] G. Cao, X. Xie, W. Yang, Q. Liao, G. Shi, and J. Wu, "Feature-fused SSD: fast detection for small objects [C]," in *Proceedings of the International Conference on Graphic and Image Processing*, p. 106151E, Qingdao, China, October 2017.

[30] Z. Zhang, G. Wen, and S. Chen, "Audible sound-based intelligent evaluation for aluminum alloy in robotic pulsed GTAW: mechanism, feature selection, and defect detection," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 2973–2983, 2018.

[31] Y. Wang, R. Guan, and Y. Lu, "Nonlinear Lamb waves for fatigue damage identification in FRP-reinforced steel plates," *Ultrasonics*, vol. 80, pp. 87–95, 2017.

[32] K. Tsukada, Y. Majima, Y. Nakamura et al., "Detection of inner cracks in thick steel plates using unsaturated AC magnetic flux leakage testing with a magnetic resistance gradiometer," *IEEE Transactions on Magnetics*, vol. 53, no. 11, pp. 1–5, 2017.

[33] S. Doshvarpassand, C. Wu, and X. Wang, "An overview of corrosion defect characterization using active infrared thermography," *Infrared Physics & Technology*, vol. 96, pp. 366–389, 2019.

[34] M. Xiao, M. Jiang, G. Li, L. Xie, and L. Y, "An evolutionary classifier for steel surface defects with small sample set [J]," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, pp. 1–13, 2017.

[35] G. Psuj, "Multi-sensor data integration using deep learning for characterization of defects in steel elements," *Sensors*, vol. 18, no. 2, p. 292, 2018.

[36] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: a survey," 2019, https://arxiv.org/abs/1905.05055.

[37] Y. Li, Y. Chen, N. Wang, and Z. -X. Zhang, "Scale-aware trident networks for object detection [J]," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6054–6063, Seoul, Korea (South), November 2019.

[38] H. Zhang, J. Xue, and K. Dana, "Deep TEN: texture encoding network [C]," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 708–717, Honolulu, HI, USA, July 2017.

[39] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, p. 1, 2017.

[40] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 405–420, Honolulu, HI, USA, 2018.

[41] G. Ren and X. Qu, "New branch optimization design based on RefineNet [C]," in *Proceedings of the IEEE International Conference on Artificial Intelligence and Computer Applications*, Dalian, China, June 2020.

Retraction

Retracted: IoT-Oriented Wireless Sensor Network and Sports Dance Movement Perception

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Zhu, "IoT-Oriented Wireless Sensor Network and Sports Dance Movement Perception," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6477170, 8 pages, 2022.

Hindawi

Research Article

IoT-Oriented Wireless Sensor Network and Sports Dance Movement Perception

Lin Zhu [ID]

College of Arts, Wuhan Sports University, Hubei, Wuhan 430079, China

Correspondence should be addressed to Lin Zhu; 2010011@whsu.edu.cn

The Internet of Things needs to connect different types of sensors in accordance with the agreement, and different agreements can be used for information and information exchange systems, which can facilitate the identification and control of intelligent systems. In many scenarios, location service application data are very important for obtaining accurate location information about nodes. In this paper, a wrist motion sensing sensor based on the PVDF piezoelectric film is developed. To realize the monitoring of wrist motion signals, this paper designs and manufactures a series of PVDF noninvasive sensors for motion perception. The characteristics of sports dance movements are reflected in the coordination between men and women, and the dance posture and movements must be synchronized with the music according to the level of music rhythm structure. The transfer of center of gravity is the main driving force of sports, and the gait of athletes is related to the transfer of center of gravity. Posture stability is the basic element of sports dance training. Only on this basis can it be possible to develop specific motion function trajectories and technical action forms.

1. Introduction

The Internet of Things has certain development prospects in some broad fields, and each positioning information needs to be accurately found. In many scenarios, location service application data are very important for obtaining accurate location information about the site [1]. Therefore, the research on precise positioning of the Internet of Things has a great social and commercial significance; for individuals, it can also effectively improve the quality of life [2]. Based on the in-depth study on the research results of the positioning algorithm of the Internet of Things, the improvement of the positioning algorithm can further improve the positioning accuracy [3]. The Internet of Things expands based on the Internet [4], connecting people, people and things, and things with things [5]. With the continuous development of the Internet of Things, sensors have been applied more and more [6]. In some applications, location information is very important, and the data used can only be effective if it contains appropriate location information. Generally speaking, the

development of a high-precision, low-complexity wireless sensor network positioning algorithm has always been an urgent need in the field of wireless research [7]. Technical problems are solved. This article takes some equipment as a starting point and then starts with the design of each unit to realize the detection of each motion signal. Investigate and study some key issues in the PVDF sensor array developed [8]. The ultimate goal of this article is to detect motion signals, use gesture recognition algorithms to realize finger motion perception, design and manufacture a set of PVDF sensor array devices, and provide a new and convenient human-computer interaction device for man-machines [9, 10]. This article does not study the hand movements because according to the sports dance technique, the standard dance arm is a fixed frame type, and different dance movements are issued by the dance arm of the body, passing through the shoulder and foot bones to the terminal fingers and then become the most basic law of movement. The limitation of the quality of sports dance technical movements comes from the results of competitive attributes, evaluation standards, and music. The

environmental limitations of sports dance technical movements come from the rules of the competition and the competition venue. As a background venue for performing technical moves, athletes can fully participate so that referees can conduct objective and rational analysis. When the quality of a technical action can be observed, it is difficult to describe its parts individually. All the athletes' technical actions must have an "obvious" effect, and any rational analysis by the referee is based on a certain "phenomenon."

2. Related Works

According to the literature, sports dance training should not only solve the kinematics and dynamics of personal technical movements at any time but also cooperate with dance partners, music, and dance style and express emotions and show themselves [11]. Under many restrictive conditions, the above elements outside the human body are deeply combined with athletes' body movements and artistic expression to show the quality of technical movements, and the pursuit of perfect technical movement quality is the competition content and winning goal [12]. The literature points out that an important technical feature of sports dance and the important feature of human independence is to rely on movement. The beginning of sports dance is reflected in the cooperation between men and women, which is mainly divided into two aspects: gait and center of gravity displacement [13]. Literature review shows that the direct research on "technical movement quality of sports dance" is extremely limited, and there is no corresponding theoretical framework to follow. The literature shows that the famous manipulator and experts in the United States have created bionic technology and can control technology in the development and design of exoskeleton robots [14]. With the rapid development of data, exoskeleton robots have developed rapidly in various countries. However, due to the gap between technology and economic level, the development of robot exoskeleton technology also varies from country to country [15]. The literature briefly introduces the current development and research results of global exoskeleton robot technology. Compared with much foreign exoskeleton technology research, the research and development of foreign skeleton robots lags behind. With the support of technology development and policy environment, action recognition technology has been greatly improved [16].

3. Internet of Things Algorithms and Wireless Sensor Network Application

3.1. Internet of Things Algorithm. Based on the positioning algorithm to measure the nodes in some positions, the initial position starts to use the weighted least square method to calculate the estimated value of the problem. Different algorithms have different functions in the positioning of the nodes in the plane.

Different nodal coordinate formulas can be obtained by three-side measurement, such as formulas (1)–(3):

$$\left(x_{it} - u_{t_1}\right)^2 + \left(y_{it} - v_{t_1}\right)^2 = d_{i,t_1}^2, \tag{1}$$

$$\left(x_{it} - u_{t_2}\right)^2 + \left(y_{it} - v_{t_2}\right)^2 = d_{i,t_2}^2, \tag{2}$$

$$\left(x_{it} - u_{t_3}\right)^2 + \left(y_{it} - v_{t_3}\right)^2 = d_{i,t_3}^2. \tag{3}$$

In formulas (4) and (5), (b) minus (a) and (c) minus (b) is the equation set as follows:

$$2\left(u_{t_1} - u_{t_2}\right)x_{it} + 2\left(v_{t_1} - v_{t_2}\right)y_{it} \\ = d_{i,t_2}^2 - d_{i,t_1}^2 + u_{t_1}^2 + v_{t_1}^2 - u_{t_2}^2 - v_{t_2}^2, \tag{4}$$

$$2\left(u_2 - u_{t_3}\right)x_{it} + 2\left(v_{t_2} - v_{t_3}\right)y_{it} \\ = d_{i,t_3}^2 - d_{i,t_2}^2 + u_{t_2}^2 + v_{t_2}^2 - u_{t_3}^2 - v_{t_3}^2. \tag{5}$$

The T-th guess of the I-th unknown node is

$$\begin{bmatrix} x_{it} \\ y_{it} \end{bmatrix} = \begin{bmatrix} 2\left(u_{t_1} - u_{t_2}\right) & 2\left(v_{t_1} - v_{t_2}\right) \\ 2\left(u_{t_2} - u_{t_3}\right) & 2\left(v_{t_1} - v_{t_2}\right) \end{bmatrix}^{-1} \\ \times \begin{bmatrix} d_{i,t_2}^2 - d_{i,t_1}^2 + u_{t_1}^2 + v_{t_1}^2 - u_{t_2}^2 - v_{t_2}^2 \\ d_{i,t_3}^2 - d_{i,t_2}^2 + u_{t_2}^2 + v_{t_2}^2 - u_{t_3}^2 - v_{t_3}^2 \end{bmatrix}. \tag{6}$$

Taking each average value as the initial value of i position nodes and then we get the estimated value as

$$\begin{bmatrix} \widehat{x}_i \\ \widehat{y}_i \end{bmatrix} = \left[\sum_{t=1}^{N} x_{it} \sum_{t=1}^{N} y_{it} \overline{NN} \right]^T. \tag{7}$$

Different positioning models will be expanded to determine the initial value of the position based on the description of the node, the distance between the nodes, and the known information measured by the TOA sorting technology, such as

$$D = Q\Delta + E. \tag{8}$$

Using the weighted least square method to determine the position deviation, we get

$$\begin{bmatrix} \Delta x_1 \\ \Delta y_1 \\ \vdots \\ \Delta x_n \\ \Delta y_n \end{bmatrix} = \left(Q^T T^{-1} Q\right)^{-1} Q^T T^{-1} D. \tag{9}$$

According to the process of the algorithm, the time complexity of the algorithm is calculated. On the basis of the Taylor formula, the three-sided complexity of each measurement method is compared, and the value of the number of different nodes is n, and each improved three-sided complexity of the edge measurement algorithm is greater than that of other algorithms. The main reason is that the

TABLE 1: Comparison of time complexity.

Algorithm	Time complexity
Trilateral calculation	$O(nm^3)$
Three-sided calculation and traditional Taylor series expansion	$O(nm^3) + O(nm)$
Three-sided calculation and Taylor series expansion of multivariate variables	$O(nm^3) + O(n^4)$

distance of the unknown node increases, which leads to the expansion of the Taylor series, in exchange for higher positioning accuracy. The comparison of time complexity is shown in Table 1.

Constructing the Fisher information matrix can get formula:

$$F = \begin{bmatrix} F_{xx} & F_{xy} \\ F_{xy}^{T} & F_{yy} \end{bmatrix}. \tag{10}$$

Equation (11) finds the F matrix that satisfies the root mean square error of the positioning result as

$$E\big((\widehat{\varphi} - \varphi)(\widehat{\varphi} - \varphi)^{T}\big) \ge \mathrm{CRLB}(\widehat{\varphi}) = \mathrm{tr}\big(F^{-1}\big). \tag{11}$$

When different algorithms are used to verify the validity of knowledge, it is necessary to use simulation technology for simulation calculation, and then determine the positioning algorithm of the trilateration in the Taylor series. The modeling parameter is an area of $100\,\mathrm{m} \times 100\,\mathrm{m}$, 20 nodes are randomly and uniformly distributed, and simulated independently for 1000 times. Formula 12 can be obtained as

$$\mathrm{error} = \sqrt{E\left(\sum_{i=1}^{n} (\widetilde{x}_i - x_i)^2 + (\widetilde{y}_i - y_i)^2\right)}. \tag{12}$$

Based on the different initial values of the positioning algorithm, it is necessary to use the least square method to estimate each different position, and then obtain each environment suitable for the two-dimensional plane.

The specific steps to solve the positioning algorithm are as follows:

① Establish the distance square matrix, as

$$D = \begin{bmatrix} 0 & d_{1.2}^2 & \cdots & d_{1.(n+m)}^2 \\ d_{1.2}^2 & 0 & \cdots & d_{2.(n+m)}^2 \\ \vdots & \vdots & \ddots & \vdots \\ d_{1.(n+m)}^2 & d_{2.(n+m)}^2 & \cdots & 0 \end{bmatrix}. \tag{13}$$

② After each different center matrix is determined, multiply, then we get

$$J = E - (n+m)^{-1}.I. \tag{14}$$

The matrix H after double centralization has the form

$$H = -\frac{1}{2}JDJ. \tag{15}$$

After the determination of different initial positions, the position of the anchor node and the distance between the two are measured using TOA sorting technology. Different Taylor set formulas create different models between multivariate expansions, and then we get

$$D = Q\Delta + E. \tag{16}$$

Equation (17) uses the least square method to solve the position deviation:

$$\begin{bmatrix} \Delta x_1 \\ \Delta y_1 \\ \vdots \\ \Delta x_n \\ \Delta y_n \end{bmatrix} = \big(Q^{T}Q\big)^{-1} Q^{T}D. \tag{17}$$

Equation (18) is the estimated location of the unknown node:

$$\begin{bmatrix} \widetilde{x}_1 \\ \widetilde{y}_1 \\ \vdots \\ \widetilde{x}_n \\ \widetilde{y}_n \end{bmatrix} = \begin{bmatrix} \widehat{x}_1 + \Delta x_1 \\ \widehat{y}_1 + \Delta y_1 \\ \vdots \\ \widehat{x}_n + \Delta x_n \\ \widehat{y}_n + \Delta y_n \end{bmatrix}. \tag{18}$$

The output mode of CRLB needs to be based on a certain algorithm in each positioning algorithm, and then meet the error according to the positioning as

$$E\big((\widehat{\varphi} - \varphi)(\widehat{\varphi} - \varphi)^{T}\big) \ge \mathrm{CRLB}(\widehat{\varphi}) = \mathrm{tr}\big(F^{-1}\big). \tag{19}$$

When the measured error is zero, the variance can be determined as an estimate of performance, and the Gaussian distribution of the root mean square error can be determined. Looking for the positioning algorithm, we obtain

$$\mathrm{error} = \sqrt{E\left(\sum_{i=1}^{n} (\widetilde{x}_i - x_i)^2 + (\widetilde{y}_i - y_i)^2\right)}. \tag{20}$$

3.2. Wireless Sensor Network Application. Because some higher-dimensional two-dimensional spaces are more intuitive, you only need to consider the two-dimensional situation. In the two-dimensional situation, the position information of the wireless sensor node is represented by a column vector of coordinate elements as

$$v_{ij} = \begin{cases} 1, & \text{measurable situation}, \\ 0, & \text{other}. \end{cases} \tag{21}$$

The distance between all pairs of nodes is measured as

$$r_{ij} = \left\| p_i - p_j \right\| + n_{ij}, i = 1, \ldots, M, j = 1, \ldots, M, i < j. \tag{22}$$

One technique that can be used to reasonably estimate its positioning performance is the AM positioning error of all detected sensor nodes as

$$\text{a.m.s.locazation error} = \frac{1}{M_{\text{loesen}}} \sum_{m \in S} E\left\{ \left\| \widehat{p_m} - p_m \right\|^2 \right\}. \tag{23}$$

According to the CRLB of AMS, formula (24) can be obtained as

$$\text{CRLB}_e = \text{TR}\left\{ I_e^{-1} \right\}. \tag{24}$$

The probability density function associated with an independent model of Gaussian measurement error with zero mean and proportional variance is

$$f_p\left(r_{ij}, i = 1, \ldots, M, j = 1, \ldots M, i < j | p_m, m = 1, \ldots, M\right). \tag{25}$$

4. Sports Dance Movement Quality and Movement Perception

4.1. Research on the Quality of Sports Dance Movements. The quality of the technical movements of sports dance is not only related to different cognitions of technical knowledge but also related to the performance of body movements. Different sports activities combine different functions in terms of concepts and principles of movement control. In training, one of the movement techniques of each athlete is to include some trajectories, timing, and movement postures related to speed, strength, and rhythm. However, it turns out that time, speed, and rhythm are all part of the content of time, and strength is related to the use of body energy. At the biological level, the understanding of movement is not only from the perspective of movement control theory, but starting from the right phase of human beings, different body movements are connected with the activities of the brain. When each athlete pursues the goal of the competition, he must constantly feel, experience, and create a state of complete or even extreme mobilization and

activation of the nervous system. Human gait can be divided into standing period and swing period. The normal gait support in dance movements requires vertical force (floor reaction force) to resist gravity, support body weight, and generate horizontal friction to resist the support surface (to make the body move towards all directions). Need to move in direction. In addition, the posture control strategy needs to be adjusted to make it more flexible, and the speed change and direction also need to be changed. In the change of each support surface, the principle of motion control will complete a certain posture control so that different speed changes can be changed. The basic principle of movement is also based on the training and application of sports dance movement. The athlete's gait model needs new adaptation, especially facing various complex tasks and environmental conditions.

As a spatial component of the quality of technical movements, the most important function of trajectory and movement methods is to control the body and prepare for the victory of the game. It can reflect the beauty more than art sports and other competitive sports. Therefore, this article examines the relationship between the quality of technical movements and the athletes and space from the perspective of kinematics, and whether the form and trajectory of the movements are related to posture positioning. Technical actions are based on body anatomy and vertical body lines and movement in spherical space. Whether it is basic movements or the movements of standing, walking, running, jumping, turning, and bending, the control of balance is a complex sensation and movement technique. Always keep this in mind during training and take the initiative to adapt to various changes. It is necessary to know that "the space where the body is located will affect the perception of position stability and direction." Posture stability is the basic element of sports dance sports training. Only on this basis can it be possible to develop specific motor function trajectories and technical movements. Amplitude, acceleration, and speed are a special technical action mode, and they are the external effect components of "mass."

4.2. The Perception of Lower Limb Movement in Sports Dance. The movement of the lower limbs of the human body contains many aspects and types of data although the use of muscle electrical signals can directly reflect the advantages of the brain's muscular awareness of the human body's actions and quickly reflect a person's movement intention. In addition, due to the uncertainty of the environment, the surface EMG signal is also more susceptible to interference. Different surface area signals are used to determine changes in motion status and ensure the accuracy and reliability of body motion. Based on the above considerations, this article assists the inertia and inertia of the lower limbs. Foot pressure information perceives movement behavior and strives to rely on the combination
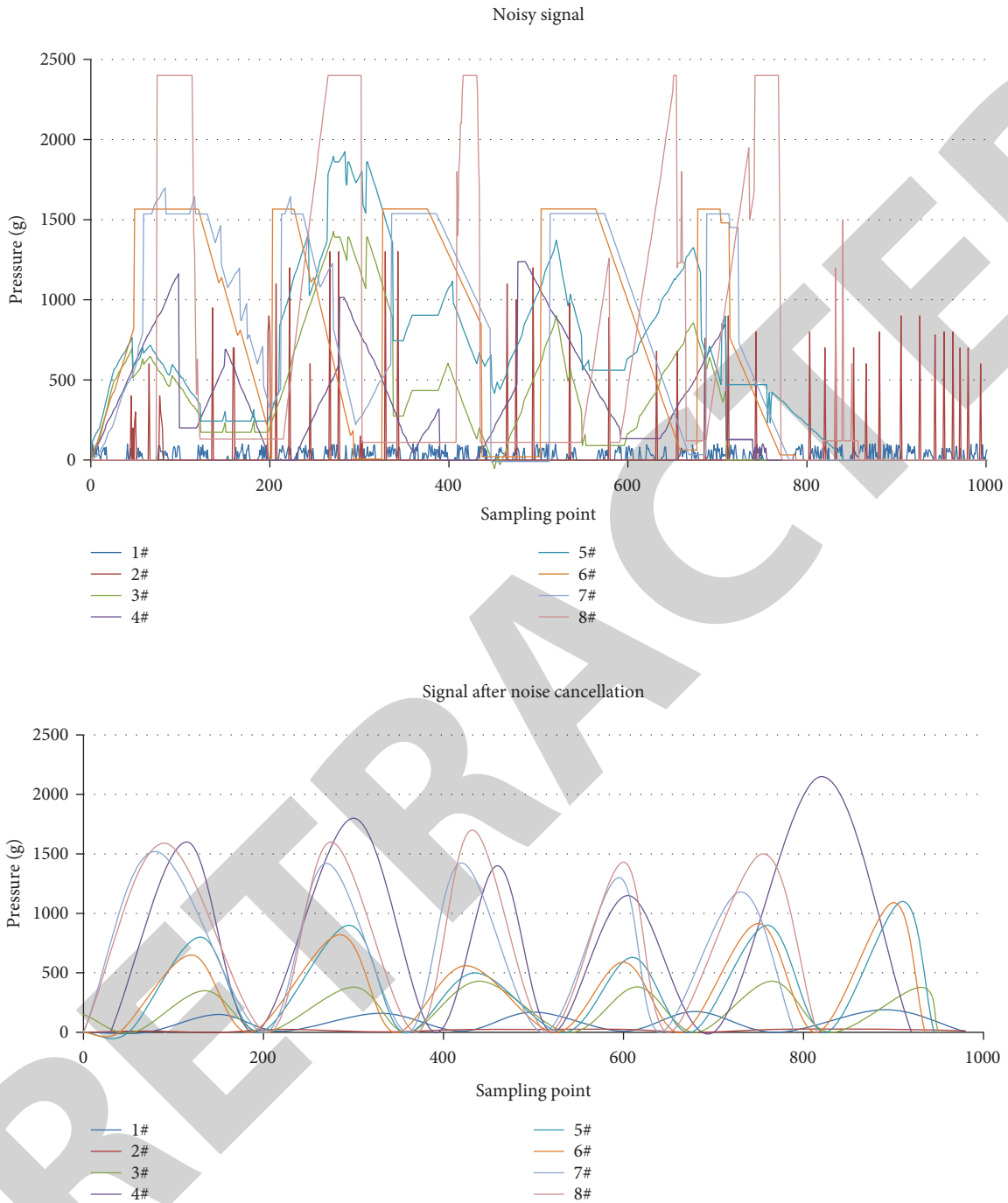
Figure 1: Low-pass filtering of plantar pressure.

of multiple sports information to make lower limb movement research more accurate and reliable, as shown in Figure 1.

Taking the lateral thigh muscles as an example, using the above-mentioned method of detecting the end point of the activity segment of the EMG signal can better identify the end point of the activity of the EMG signal, as shown in Figure 2.

After recording the acceleration characteristics, sEMG, and plantar pressure signals separately, some important characteristics can be obtained, but this also leads to an increase in the size of the elements, and the excess elements
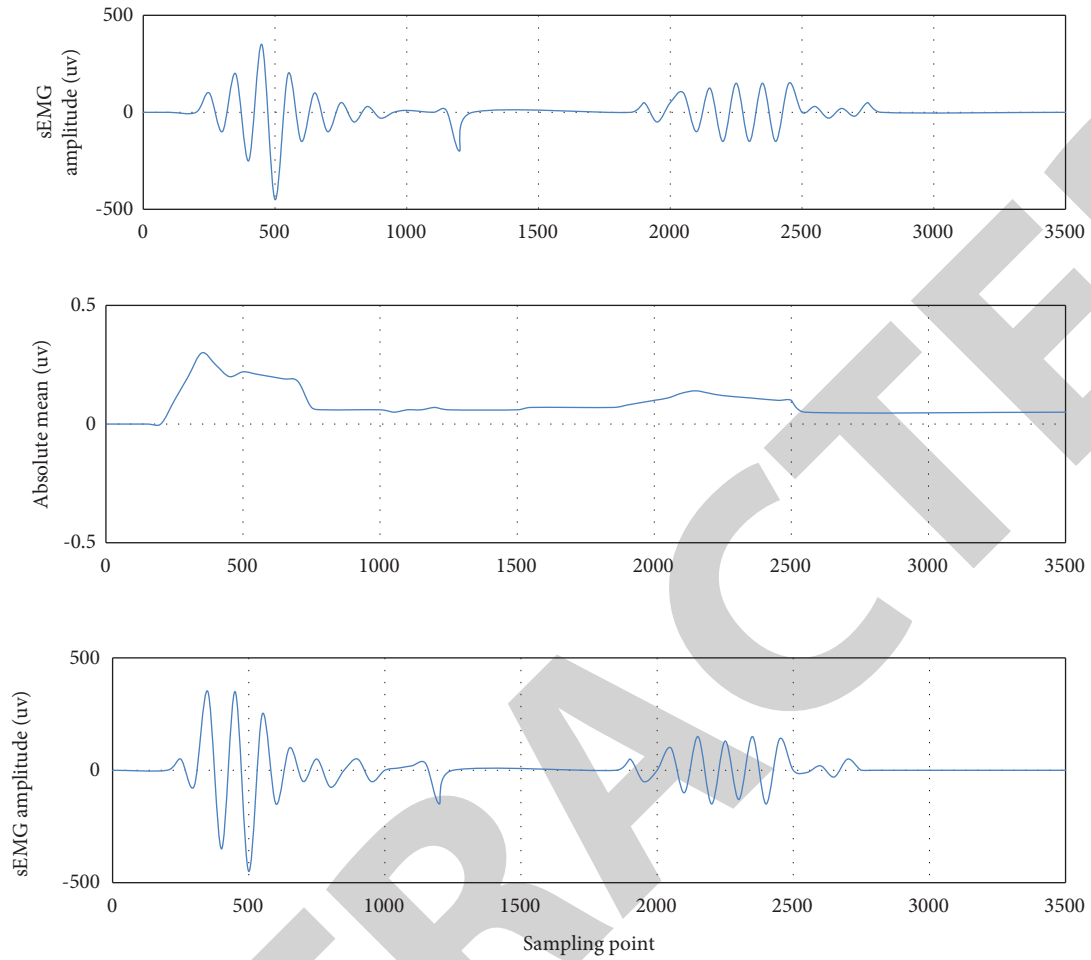
FIGURE 2: Schematic diagram of active segment extraction.

are also different. Therefore, the preliminary merge function should be checked. This article uses a method to check the relevance and redundancy of the function, as shown in Figure 3.

Table 2 shows the comparison and analysis results of the classification effect data. From the results obtained, the classification accuracy of the algorithm combining the three types of information functions is compared with the method of a single sensor function and compared with the previous method using the gait periodic function as the unit. Compared with the recognition method, the classification accuracy is higher, and at the same time, it has a great advantage in classification efficiency.
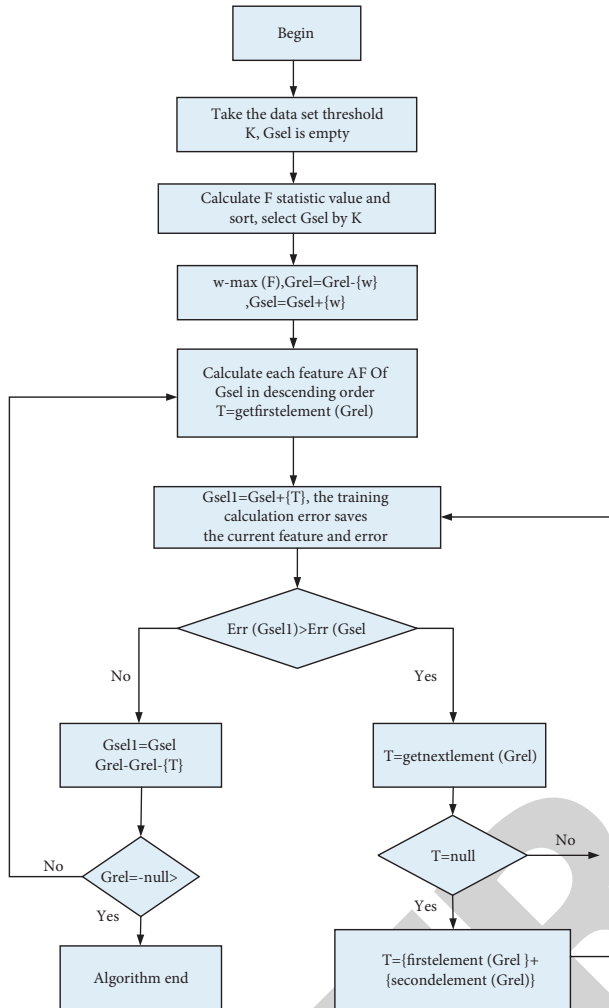
FIGURE 3: Flow chart of the forward search algorithm based on dynamic correlation and redundancy analysis.

TABLE 2: Comparison of three feature classification results.

	Walk (%)	Jump (%)	Raise leg (%)	Stretch (%)
Original period unit	90.12	89.10	92.50	89.98
sEMG	81.32	82.50	89.00	81.20
Acceleration	87.30	87.20	83.92	86.17
Plantar pressure	90.35	91.50	78.65	77.15
Fusion feature	88.93	93.80	91.32	91.56

5. Conclusion

This article aims to further reduce the influence of ranging error on positioning accuracy and improve the positioning accuracy of the algorithm. Based on the Taylor series expansion method, in each of the expansion positioning models, which are mostly Taylor series, each model proposes two brand new positioning algorithms, which combine the multilateral measurement methods, and then uses the MATLAB algorithm to model and test the algorithm. In order to be able to model and analyze the positioning algorithm more intuitively and conveniently, this paper also uses MATLAB to design and implement an IoT node positioning modeling platform. Different environmental monitoring has attracted widespread attention in areas such as intelligent transportation and technological battlefields. Different industrial production and life have gradually expanded to social life. However, as a continuously developing technology industry, it is still in application. As the core technology of WSNS, WSNS positioning technology has always been a research topic for scientific research institutes and enterprises at home and abroad. Different positioning algorithms emerge in endlessly, but no one positioning algorithm is omnipotent and suitable for all situations. Each algorithm has its own advantages and limitations in a given environment. Based on the analysis of the principles and performance of various positioning algorithms, this article assumes the basic requirements of low computing power and high practical feasibility of the product to develop a system that uses a small number of reference nodes. The practice of multiple scenarios requires in-depth research and positioning. Technology and algorithm goals apply part of the positioning information to the real wireless temperature monitoring system. It is the internal and external sum of technical knowledge and performance behavior concepts. It includes many factors, and only when the requirements of these factors are fully met as much as possible will it affect the quality of technical actions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that they have no conflicts of interest.

References

[1] S. Li, L. D. Xu, and S. Zhao, "The internet of things: a survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.

[2] K. Rose, S. Eldridge, and L. Chapin, "The internet of things: an overview," *The internet society (ISOC)*, vol. 80, pp. 1–50, 2015.

[3] L. D. Xu, W. He, and S. Li, "Internet of things in industries: a survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.

[4] F. Wortmann and K. Flüchter, "Internet of things," *Business & Information Systems Engineering*, vol. 57, no. 3, pp. 221–224, 2015.

[5] m. Ayandele and O. S. Adeoye, "Changing work environment through information and communication technology (ICT): challenges to secretarial staff," *International Journal of Computer Application*, vol. 9, no. 10, pp. 35–40, 2010.

[6] L. Zheng, W. Zhou, W. Tang, X. Zheng, A. Peng, and H. Zheng, "A 3D indoor positioning system based on low-cost

Hindawi

Research Article

Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIoMT Approach

Mohanad A. Deif [1], Hani Attar [2], Ayman Amer,[2] Ismail A. Elhaty [3], Mohammad R. Khosravi [4], and Ahmed A. A. Solyman [5]

[1]*Department of Bioelectronics, Modern University for Technology and Information (MTI), Cairo, Egypt*
[2]*Department of Energy Engineering, Zarqa University, Zarqa, Jordan*
[3]*Department of Nutrition and Dietetics, Faculty of Health Sciences, Istanbul Gelisim University, Istanbul, Turkey*
[4]*Department of Computer Engineering, Persian Gulf Engineering, Bushehr, Iran*
[5]*Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Nişantaşi University, 34398, Istanbul, Turkey*

Correspondence should be addressed to Mohammad R. Khosravi; m.khosravi@mehr.pgu.ac.ir

Overall prediction of oral cavity squamous cell carcinoma (OCSCC) remains inadequate, as more than half of patients with oral cavity cancer are detected at later stages. It is generally accepted that the differential diagnosis of OCSCC is usually difficult and requires expertise and experience. Diagnosis from biopsy tissue is a complex process, and it is slow, costly, and prone to human error. To overcome these problems, a computer-aided diagnosis (CAD) approach was proposed in this work. A dataset comprising two categories, normal epithelium of the oral cavity (NEOR) and squamous cell carcinoma of the oral cavity (OSCC), was used. Feature extraction was performed from this dataset using four deep learning (DL) models (VGG16, AlexNet, ResNet50, and Inception V3) to realize artificial intelligence of medial things (AIoMT). Binary Particle Swarm Optimization (BPSO) was used to select the best features. The effects of Reinhard stain normalization on performance were also investigated. After the best features were extracted and selected, they were classified using the XGBoost. The best classification accuracy of 96.3% was obtained when using Inception V3 with BPSO. This approach significantly contributes to improving the diagnostic efficiency of OCSCC patients using histopathological images while reducing diagnostic costs.

1. Introduction

Oral squamous cell carcinoma (OSCC) is a diverse collection of cancers that arise from the mucosal lining of the oral cavity [1, 2], accounting for more than 90% of all oral cancers [3]. It is a subtype of head and neck squamous cell carcinoma (HNSCC), which is the world's seventh most frequent cancer [4]. The World Health Organization estimates that 657,000 fresh cases are diagnosed annually, with over 330,000 fatalities globally. OSCC frequency rates were found to be massively greater in South Asian countries. India has the biggest number of cases (one-third of cases), whereas

Pakistan has the first and second most common cancers in males and females, respectively [5]. Drinking alcohol, smoking cigarettes, poor oral hygiene, human papillomavirus (HPV) exposure, genetic background, lifestyle, ethnicity, and geographical region are all risk factors.

The detection of OSCC in early stages is essential to achieve a successful therapy, increased chances for survival, and low mortality and morbidity rates [6]. With a 50% average cure rate, the OSCC has bad prognosis [7, 8]. Microscopy-based histopathological analysis of tissue samples is considered the standard method for diagnosing OSCC [9, 10]. This diagnostic pathology methodology depends on

histopathologists' interpretation, which is typically slow and error-prone, limiting its clinical utility [11]. As a result, it is critical to provide effective diagnostic tools to aid pathologists in the assessment and diagnosis of OSCC.

Lately, there have been a growing number of studies on applying artificial intelligence (AI) to improve medical diagnostics. Researchers have been able to examine AI applications in medical image analysis thanks to increasing the usage of diagnostic imaging. Deep learning (DL) [12], in particular, has shown outstanding success in solving a variety of medical image processing challenges [13], specifically in the diagnostics of pathological images [14, 15]. Computer-aided diagnosis (CAD) systems based on DL have been suggested and established on a large scale for a range of cancer sorts, such as breast cancer, prostate cancer, and lung cancer ([16, 17] and [18]). Nevertheless, the research shows that DL has been seldom used to diagnose OSCC from pathological images. To recognize keratin pearls in oral histopathology images, Dev et al. employed Convolutional Neural Network (CNN) and Random Forest. For keratin area segmentation, the CNN model achieved 98.05 percent accuracy, whereas the Random Forest model spotted keratin pearls with 96.88 percent accuracy [19]. Das et al. used DL to divide oral biopsy images into different classes according to Broder's histological grading system. CNN was also proposed, which had a 97.5 percent accuracy rate (N) [20].

Folmsbee et al. used CNN to classify oral cancer tissue into seven types using Active Learning (AL) and Random Learning (RL) (stroma, tumor, lymphocytes, mucosa, keratin pearls, adipose, and blood). The AL's accuracy was determined to be 3.26 percent higher than that of the RL [21]. Furthermore, Martino et al. applied multiple DL architectures, such as U-Net, SegNet, U-Net with VGG16 encoder, and U-Net with ResNet50 encoder, to segment oral lesion whole slide images (WSI) into three groups (carcinoma, noncarcinoma, and nontissue). A deeper network, such as U-Net upgraded with ResNet50 as encoder, was shown to be more accurate than the original U-Net [22]. Amin et al. recently used VGG16, Inception V3, and Resnet50 were fine-tuned individually and then used in concatenation as a feature extractor to perform binary classification on oral pathology images [23].

The categorization of oral histopathological images into normal and OSCC classes was enhanced in this research. The increase is achieved by employing the notion of transfer learning to extract features from pretrained CNN models (VGG16, AlexNet, ResNet50, and Inception V3). Then, using the Binary Particle Swarm Optimization (BPSO) approach, features were selected. Finally, three different classifiers, XGBoost, KNN, and ANN, were used to determine the detection performance of OSCC and normal oral histopathology images.

2. Research Contributions

The proposed research investigates the following:

   (a) Oral histopathological images introduced to extract the desired features to four various models of DL.

TABLE 1: Image categories with the related source.

Resolution	Class	Number of images
100× magnification	NEOR	89
	OSCC	439
400× magnification	NEOR	201
	OSCC	495
Total images	NEOR and OSCC	1224

   (b) Investigating the obtained features in "a" by applying the methods of metaheuristic feature selection.

   (c) The metaheuristic feature selection method is explored to enhance the performance of the classifiers.

   (d) Achieving the most recommended "hybrid" system that provides the highest accuracy in the stage of the making decision, which is obtained by comparing the highest performance of the proposed models for detecting OSCC.

   (e) The final stage of the proposed research is the effect evaluation for the stain normalization on images classification.

   (f) Based on "a" to "e," AIoMT platforms development [24–26] is regarded as the main goal of the introduced research in the medical and health field.

The rest of the paper is organized as follows: The literature study in Section 1 goes into detail on previous successful approaches. A discussion of the method and approaches used in this study can be found in Section 3. A breakdown of the findings from the study is given in Section 4. Finally, the outcome of the investigated study is described in Section 5 before concluding the paper in Section 6.

3. Materials and Methods

3.1. Materials. The histopathological images dataset is got from the online resource repository [27], and medical experts prepared and cataloged data from 230 individuals. These images were collected from two different locations in India: Dr. B. Borooah Cancer Research Institute and Ayursundra Healthcare Pvt. Ltd, Guwahati. Images are divided into two categories: normal epithelium of the oral cavity (NEOR) and OSCC. A Leica ICC50 HD microscope camera was used to get the images. Table 1 shows the histopathological image details in terms of class, resolutions, and quantity. Figure 1 shows a chosen selection of histopathological images and the related categories.

3.2. Proposed Model. Figure 2 depicts the proposed model. The model states that the input histopathology images were first preprocessed. In order to extract features from normalized histopathology images, four DL models were used. The transfer learning approach was used during the feature extraction procedure. To find the best probable features, Binary Particle Swarm Optimization was employed as a feature selection method. Finally, three machine learning

(a)

(b)
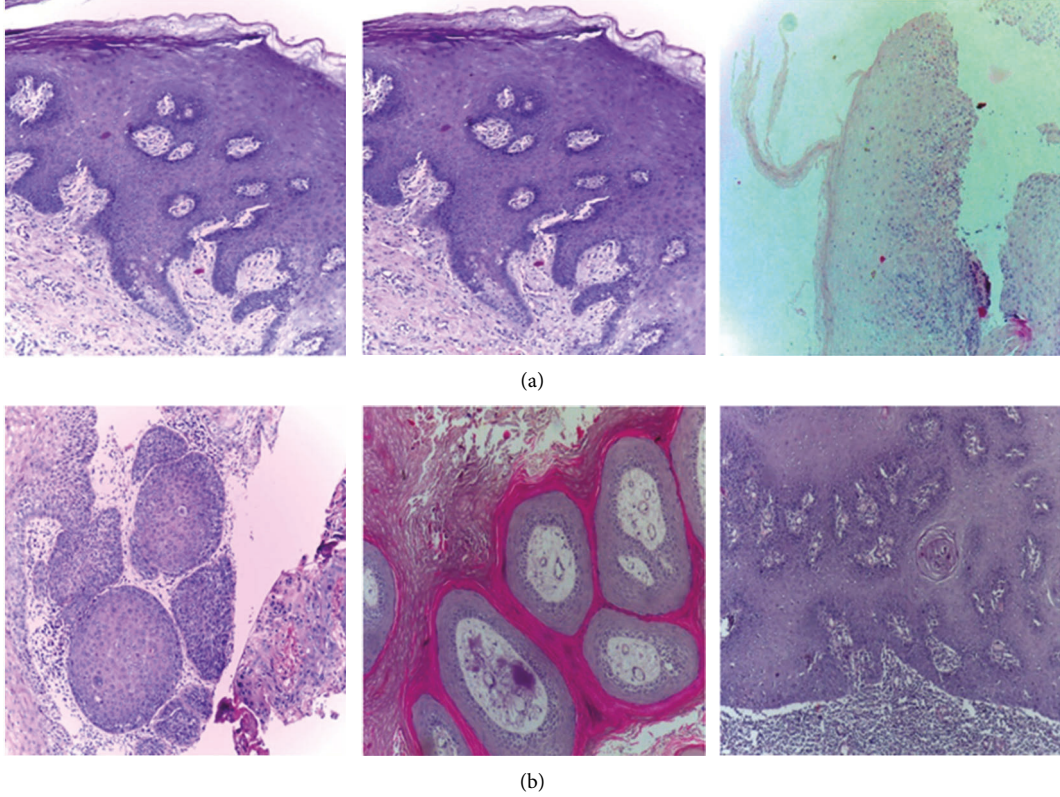
FIGURE 1: Different classes of histopathological images with (a) NEOR and (b) OSCC.

methods were used to classify abstract features, such as DL [28].

3.2.1. Image Enhancement.

To develop the classification process, some preprocessing work is regarded as essential. So, the processes performed in the preprocessing step are explained in the following.

(1) Image Resizing. As shown in Table 2, the size of the histopathology images was adjusted depending on the size of the input of each DL architecture. Losing local features in histopathological images when resized is offset by the preservation of global features.

(2) Image Normalization. The analysis of the histopathology images faces may challenges, such as producing robust models to the variations resulting from different labs and imaging systems [30], caused by raw materials, the response to various colors of the scanners' slide, protocols for staining, and manufacturing techniques [31]. The histopathology dataset was stain-normalized using the Reinhard method [32, 33]. The purpose of the Reinhard stain normalization approach is to bring these stains' appearances into line [34]. This method allows a target image to be shifted in the color domain to more closely match the template image, which is applied on each histopathology image pixel as shown in equations 1 to 3 [35, 36].

Step 1: Convert both the source image X and target image Y from RGB space to $l\alpha\beta$ space.

Step 2: Do the following transformation in $l\alpha\beta$ space:

$$l_2 = \mu_g(l_1) + [l - \mu_g(l)] * \left[\frac{\sigma_g(l_1)}{\sigma_g(l)}\right], \quad (1)$$

$$a_2 = \mu_g(a_1) + [a - \mu_g(a)] * \left[\frac{\sigma_g(a_1)}{\sigma_g(a)}\right], \quad (2)$$

$$b_2 = \mu_g(b_1) + [b - \mu_g(b)] * \left[\frac{\sigma_g(b_1)}{\sigma_g(b)}\right], \quad (3)$$

where $l_2, a_2,$ and b_2 are intensity variables of the processed image in $l\alpha\beta$ space. $l_1, a_1,$ and b_1 are intensity variables of target image in $l\alpha\beta$ space. $l, a,$ and b are intensity variables of source image. μ_g indicates the global mean of the image and σ_g represents the global standard deviation of the image.

Step 3: Convert back the processed image Z from $l\alpha\beta$ space to RGB space.

The intensity variation of the original image is preserved using this procedure. As a result, its structure is kept, while the contrast is altered to match that of the target.

3.2.2. Features Extraction Using Deep Learning Models.
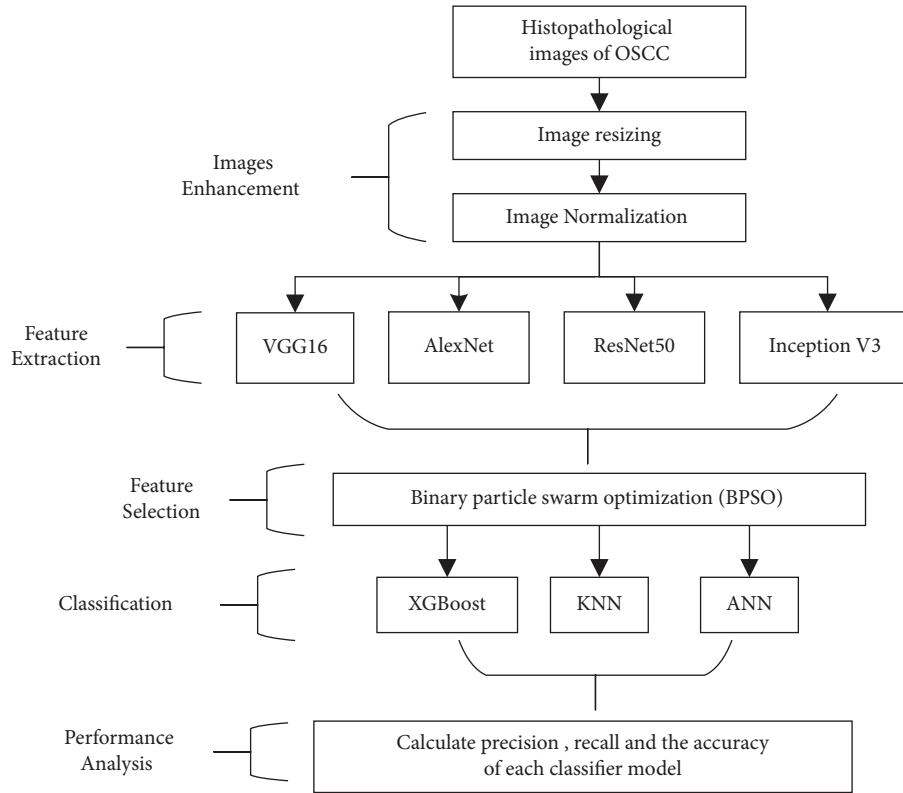
Figure 2: The proposed model for realizing AIoMT in medical systems.

Table 2: The input size for each DL architecture [29].

Model	Input size (pixels)
VGG16	224×224
AlexNet	227×227
ResNet50	224×224
Inception V3	299×299

In this work, four DL models were used for feature extraction, namely, VGG16, AlexNet, Inception V3, and ResNet50.

*(1) VGG*16. Maximum pooling and the convolution layers arrangement are followed by the VGG16 architecture [37] (Figure 3). Eventually, there are three FC layers: the first and the second are triggered by ReLU, while the third is triggered by Softmax. This design has 16 layers and 138 million parameters; the input layer can receive images with 224×224 pixels.

(2) AlexNet. The AlexNet [38] model contains 61 million parameters using an 8-layer CNN architecture. In the AlexNet architecture (Figure 4), there are five convolutional layers, where three of them are only used as linked layers. The fourth layer is the Softmax layer that requests a resolution of 227×227 pixels input image, and the last layer is the ReLU activation function, which performs the system convolutional and connected processes. Moreover, the FC-8 layer is linked to the Softmax layer via 39 neurons.

*(3) ResNet*50. The ResNet50 architecture [39] addressed the issues of several nonlinear layers, not learning identity mappings, and deterioration. Simply said, ResNet50 is a network made up of residual unit stacks (Figure 5). The network is built using residual units as building components. These units were built with convolution and pooling layers. The input histopathology images have a resolution of 224×224 pixels, and the design includes 3×3 filters.

*(4) Inception V*3. Convolutions, average pooling, maximum pooling, dropouts, and completely connected layers are among the asymmetrical and symmetrical building elements used in the model (Figure 6). The Softmax function is found in the last layer of the Inception V3 architecture [40], which comprises 42 levels, where the resolution of the received information by the input layer in pixels is 299×299.

3.2.3. Binary Particle Swarm Optimization for Feature Selection. In processing data, the selection of features is critical [41–43]. Longer training and overcompliance are challenges caused by the vast amount of data to be handled (Deif et al. [12, 44]). Unnecessary features should be removed from the data to avoid situations like this. For feature selection [7], BPSO is used, where BPSO is the binary form of PSO [45, 46]. The BPSO algorithm process was summarized in Figure 7 and described below:

> Step (1): Equation (4) is used to update the particle velocity in the swarm.
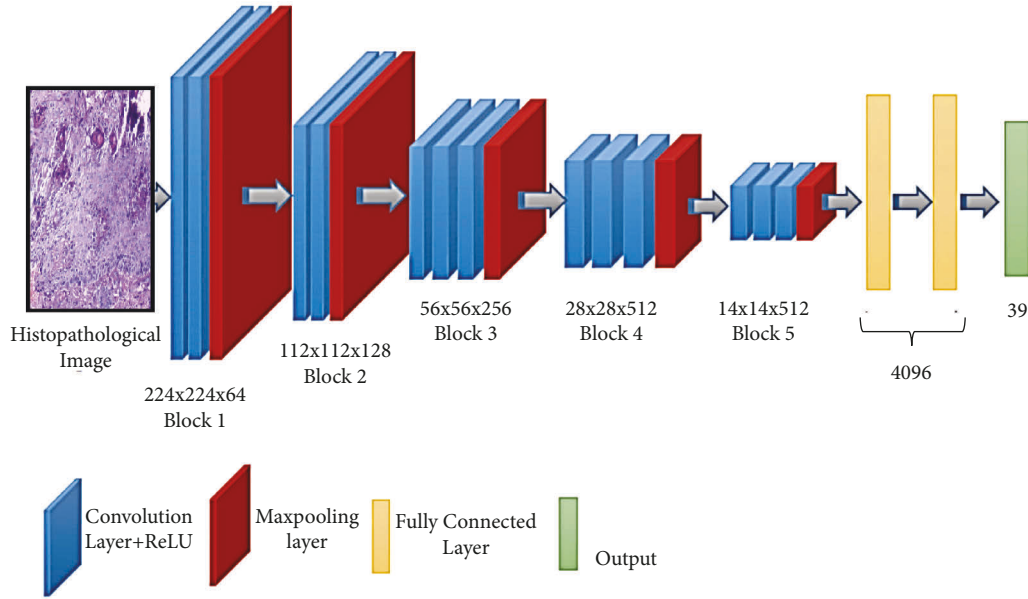
Figure 3: Schematic illustration of VGG16 architecture.
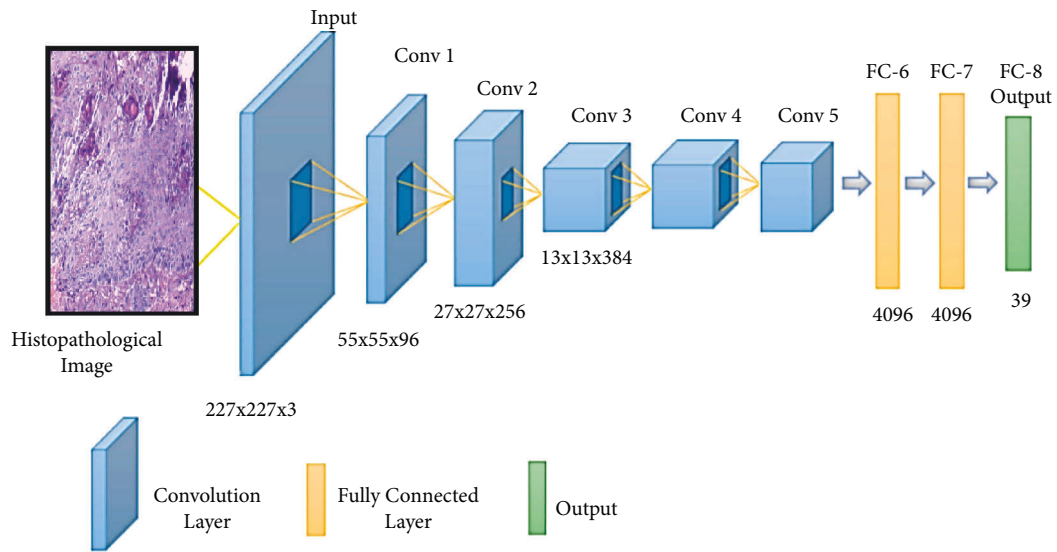
Figure 4: Schematic illustration of AlexNet architecture.
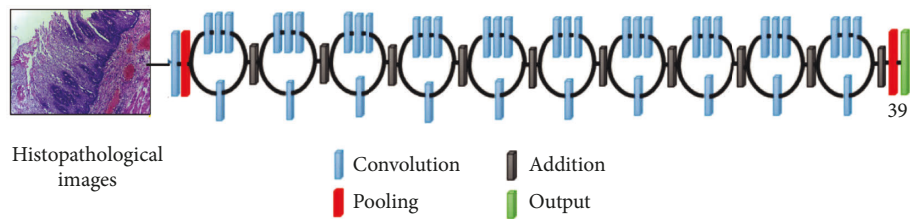
Figure 5: Schematic illustration of ResNet50 architecture.
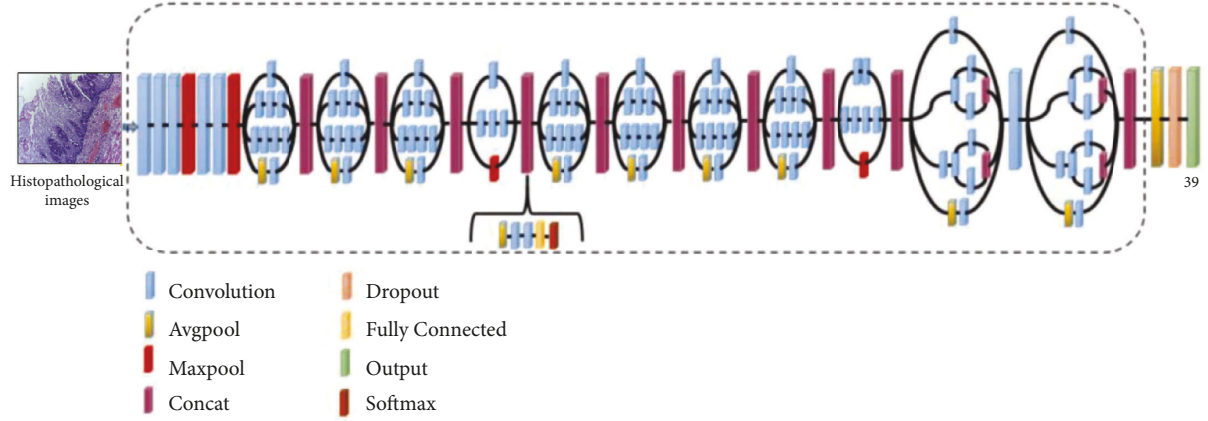
FIGURE 6: Inception V3 architecture schematic.
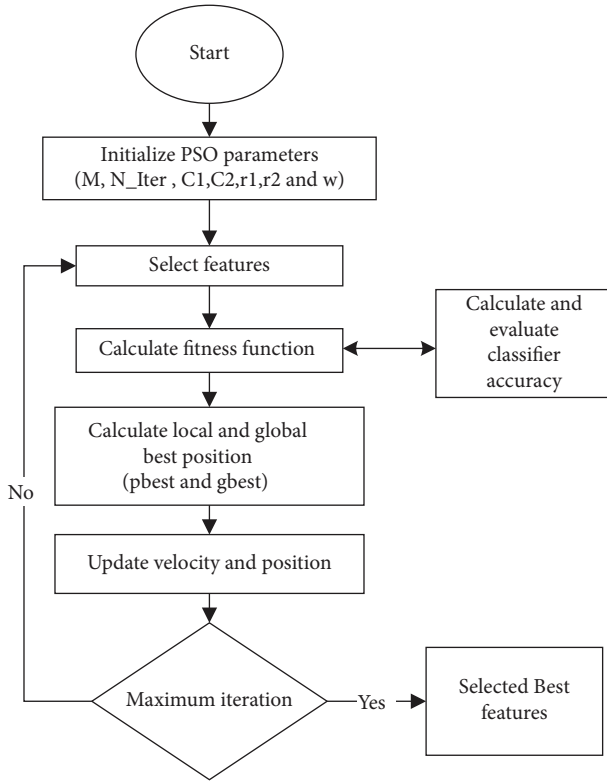
FIGURE 7: BPSO approach for feature selection.

$$v_i^d(t+1) = wv_i^d(t) + c_1 r_1 \left(pbest_i^d(t) - x_i^d(t) \right)$$
$$+ c_2 r_2 \left(pbest^d(t) - x_i^d(t) \right). \tag{4}$$

Step (2): Equations (5) and (6) are used to modify the particle's location.

$$S\left(v_i^d(t+1)\right) = \frac{1}{1 + \exp\left(-v_i^d(t+1)\right)}, \tag{5}$$

$$x_i^d(t+1) = \begin{cases} 1, & \text{if rand} < S\left(v_i^d(t+1)\right). \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Each swarm particle represents a potential solution. A 0/1 (1x feature number) matrix is one conceivable option, with each column value produced at random. Selecting the column numbers in the feature matrix with 1 yields a workable solution (a particle). The fitness value of particles in a swarm was calculated using the kNN classifier error rate. The particle with the best objective function value in the swarm is considered as g_{best} in the first iteration.

Step (3): The particle with the highest aim function value is allocated in subsequent iterations (equation (7)).

$$Pbest \begin{cases} id(t+1) = xi(t+1), & if F\left(x_i(t+1)\right) < F\left(pbest_i(t)\right). \\ pbest_i(t), & \text{otherwise.} \end{cases} \tag{7}$$

If p_{best}'s aim function value is higher than g_{best}'s, p_{best} is assigned to G_{best} (equation (8)).

$$gbest(t+1) = \begin{cases} pbest_i, (t+1), & if F\left(pbest_i(t+1)\right) < F\left(gbest(t)\right), \\ gbest(t), & \text{otherwise,} \end{cases} \tag{8}$$

where x is the solution, p_{best} denotes the personal best, B_{best} denotes the global best, $F(.)$ denotes the fitness function, and t denotes the number of iterations.

This process is repeated for T iterations, and the algorithm returns the best global solution as output [7].

3.2.4. Classification Stage. Xtreme gradient boosting (XGBoost) achieves the classification model for distinguishing between histopathological images for NEOR and OSCC. The findings of XGBoost are compared to those of classic machine learning techniques like artificial neural networks (ANN) and Random Forest (RF) [47].

3.2.5. Assessment of the Suggested Methodology's Classification Performance. The suggested methodology is evaluated

(a)   (b)   (c)
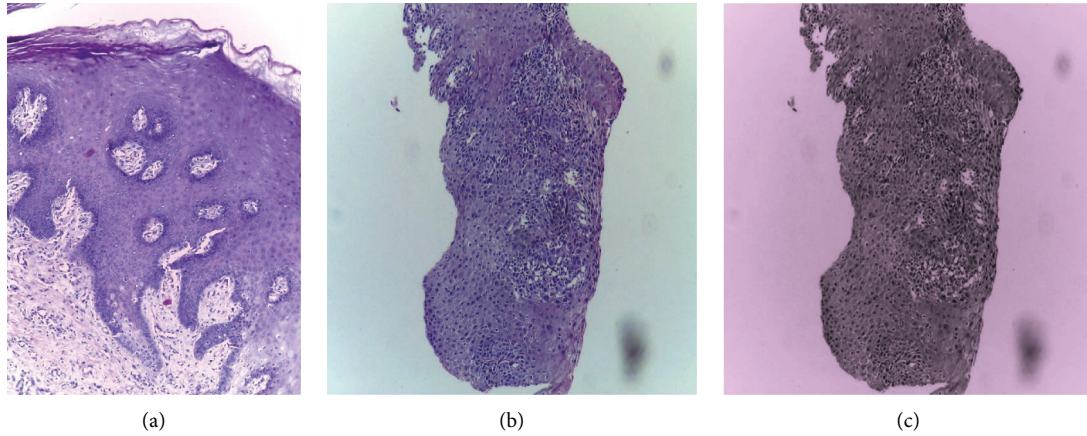
Figure 8: Samples of original and normalized histopathological images: (a) original image; (b) target image; and (c) Reinhard normalized.

by calculating sensitivity, precision, and accuracy for classifiers. The following formula is used to calculate these terms:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\%,$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\%, \qquad (9)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%.$$

The number of correctly classified chest disease images in the normal and OSCC classes, respectively, is represented by TP (True Positive) and TN (True Negative). The numbers FN (False Negative) and FP (False Positive) in the normal and OSCC classes, respectively, show the number of misclassified histopathological images.

4. Experimental Setup and Setting

The layers of the all deep models would have been trainable true during the training phase. Softmax was used as the activation function of the final CNN layer, and MSE was used as a loss function. The early halting approach was employed, with patience set at 5 and the minimal loss change set to 0.0001. The SGD optimizer was employed in all deep models with a learning rate of 0.001. The size of the mini-batch is 32.

Histopathological images were divided into two groups in the classification stage: 80 percent for training the classifier models and 20 percent for testing, with 10-fold cross-validation. All experiments were run on Google Colab [48] with GPU support. The whole code was written in Python 3.10.1 [49] using Keras version 2.7.0 [50].

The algorithms implemented in this research are BPSO, XGBoost, ANN, and RF, taking into consideration the fact that these algorithms are varying in the hyperparameters, and the settings are shown in [51].

5. Results

In the first step, to better understand the impact of stain normalization approaches on classification performance, all

histopathology images in the dataset were stain-normalized using the Reinhard technique. The Reinhard stain normalization approach used in this study is illustrated in Figure 8. In this diagram, (a) represents the original histopathological image used for the Reinhard technique, (b) represents the target image to be transferred (the techniques aim to normalize the colors in the target images to those of the original), and (c) represents the result of using the Reinhard technique.

In addition, we employed the probability density function (PDF) to investigate the effect of the Reinhard normalization method on histopathological images by comparing the postnormalization image PDF with the prenormalization image PDF. Figure 9(a) is the original PDF. The resultant density functions of the pixels in 3-channel RGB show that the target image's distribution is substantially skewed to the left (Figure 9(b)). Figure 9(c) shows that, after normalization, the probability distribution of the three channels resembles that of the original image.

Then, four deep models were used to extract features from the images: VGG16, AlexNet, ResNet50, and Inception V3. Figure 10 shows the total number of features got from each deep model. It can be seen in Table 3 that the smallest number of extract features was from the AlexNet model and numerous extract features were got from the ResNet50 model. Meanwhile VGG16 and Inception V3 extracted the same number of features.

Two experiments were carried out to see if the proposed feature selection strategy could increase the classification model's accuracy. In the first experiment, to train the classifier, all features were used without applying the feature selection technique. The second experiment was performed by employing the proposed feature selection method, that is, BPSO, to eliminate the undesired redundant and unrelated features by looking for the best features used to train the classifier. The performance values obtained for the XGBoost classifier using all extracted features and selected features are given in Figure 11.

The findings of the experiments show that when using the selected technique on features extracted using VGG16 and Inception V3, the classification performance of the XGBoost model was reduced but not when applying the

(a)

(b)

(c)
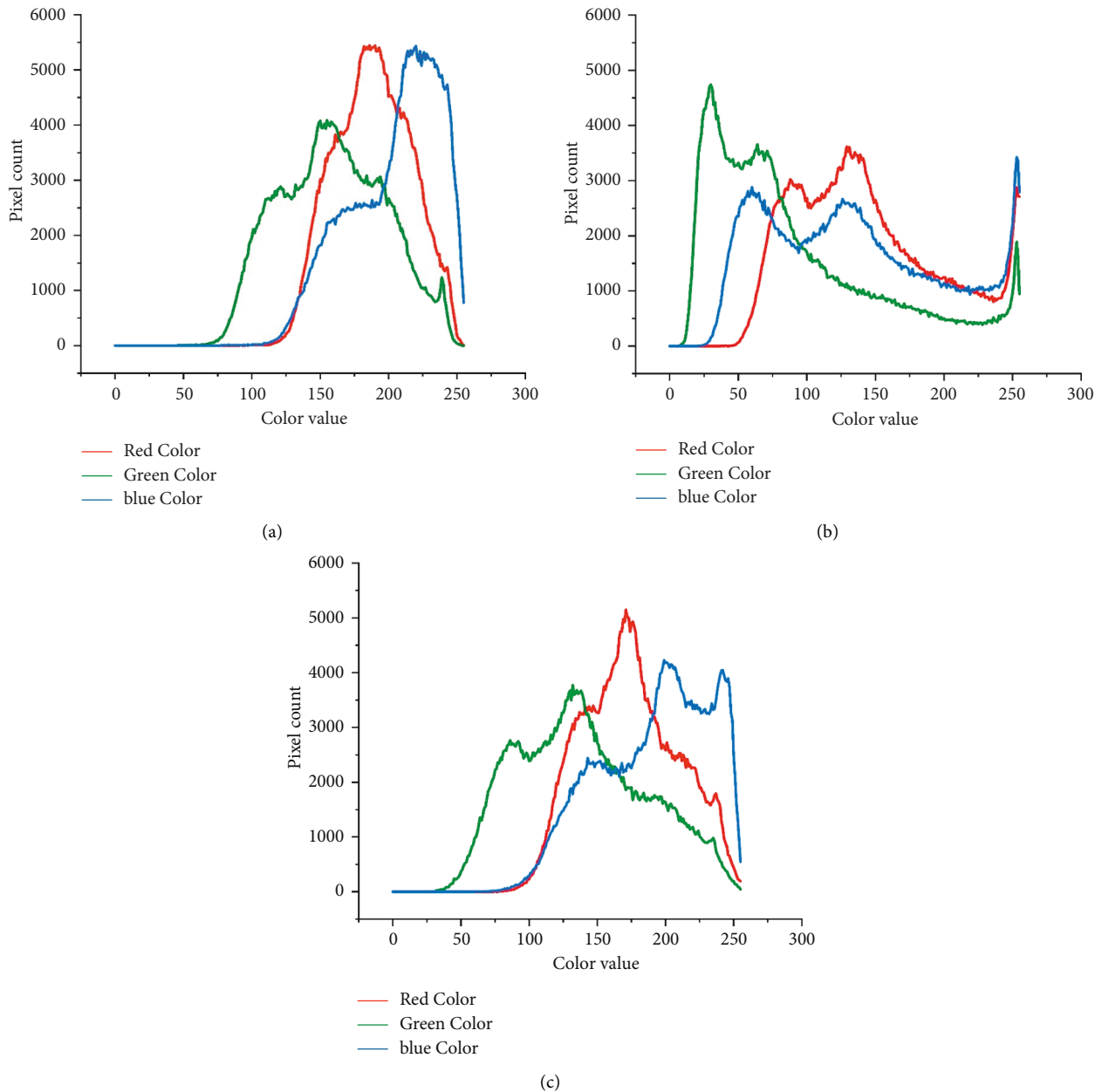
FIGURE 9: Samples of the probability density function (PDF) for (a) original histopathological image, (b) target histopathological image, and (c) Reinhard normalized.

selected approach on features derived using AlexNet. Meanwhile the XGBoost model delivered outstanding results when using the selected feature approach on features that were extracted using ResNet50.

To interpret the performance of XGBoost when using the features selection technique, Figure 12 was developed to illustrate the relation between the numbers of extracted features from all deep models; the number of features was selected using BPSO and Delta value (δ) that indicates difference between classification accuracies for XGBoost before and after applying BPSO.

It is seen that the highest extracted number belongs to the ResNet50 model, while VGG16 and Inception V3 have about half extracted number compared to the ResNet50 model. The lowest extracted number was got from AlexNet.

It is concluded that there is a relationship between the number of features extracted and the feature selection approach, where numerous features increase the ability of BPSO to select the best features and then select from fewer features. Therefore, we can interpret that the accuracy of XGBoost increased ($\delta = 11\%$) when applying BPSO on features that were extracted from ResNet50 because the BPSO selects the best features from numerous features (4099) compared to features extracted from VGG16 (2051) and Inception V3 (2069), while the XGBoost accuracy has not changed ($\delta = 0$).
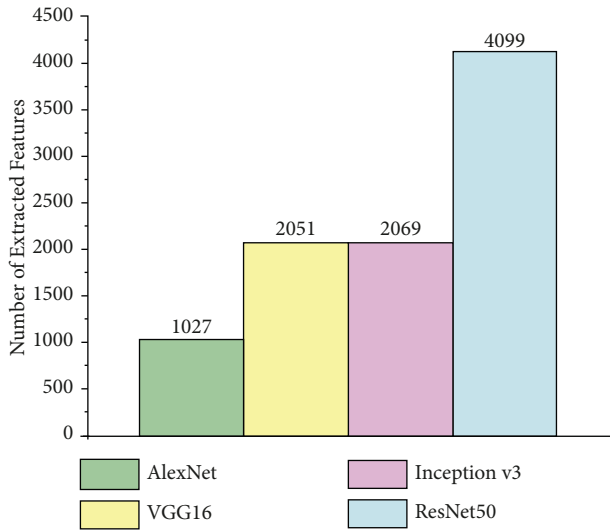
FIGURE 10: The features got in the study and all execution times.

FIGURE 12: Comparison between the number of extracted features from all deep models, number of selected features, and Delta value (δ).

TABLE 3: Performance comparison for XGBoost classifier with another state-of-the-art traditional machine learning algorithm.

Classifier name	Overall accuracy (%)	Sensitivity (%)	Precision (%)
XGBoost	96.3	98.9	96.3
RF	93.1	97.8	93.3
ANN	94.1	97.8	94.7

FIGURE 13: Fitness values of BPSO for different extracted features.
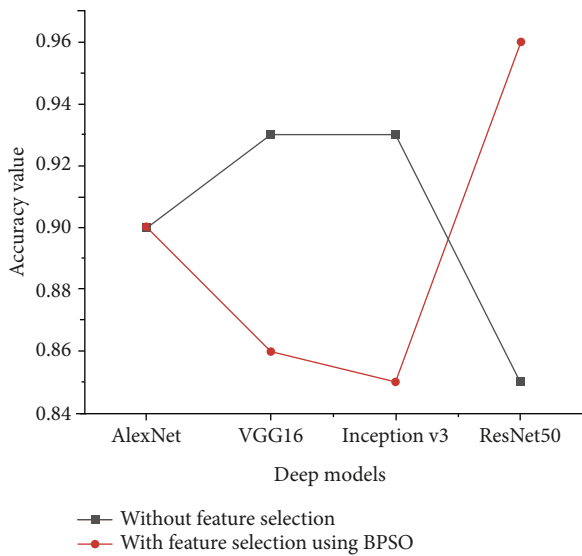
FIGURE 11: Accuracy performance for the XGBoost classifier using all extracted features and selected features.

Figure 13 illustrates a comparison of the BPSO convergence curves algorithm at the feature selection stage. It is important to note that the fitness is the average of 20 runs. The greater the performance, the lower the values of the best fitness, worst fitness, and mean fitness. It is observed that the features obtained from VGG16 and Inception V3 models achieved the lowest fitness value for BPSO, while the best performance is given by the feature group obtained by using ResNet50.

Because the XGBoost classifier's histopathological image classification results with each features extraction method are the average accuracy got from 20 independent runs, a two-sample t-test with a 95 percent confidence level was used to see if the classification performance of the proposed ResNet50 with BPSO was significantly better (p value <0.05) than those of the other methods on histopathological images.

To assess the XGBoost classifier's classification performance, artificial neural network (ANN) and Random Forest are two more traditional machine learning techniques that are compared to the outcomes (RF). The classifier models were trained using features taken from the ResNet50 model and then fed to the classifier by best-selected features using BPSO, based on previous results. A confusion matrix for all classifiers is shown in Figure 14.

It is clear from the confusion matrix that the highest TP results were got from the XGBoost classifier, where 184 of

(a)

(b)

(c)
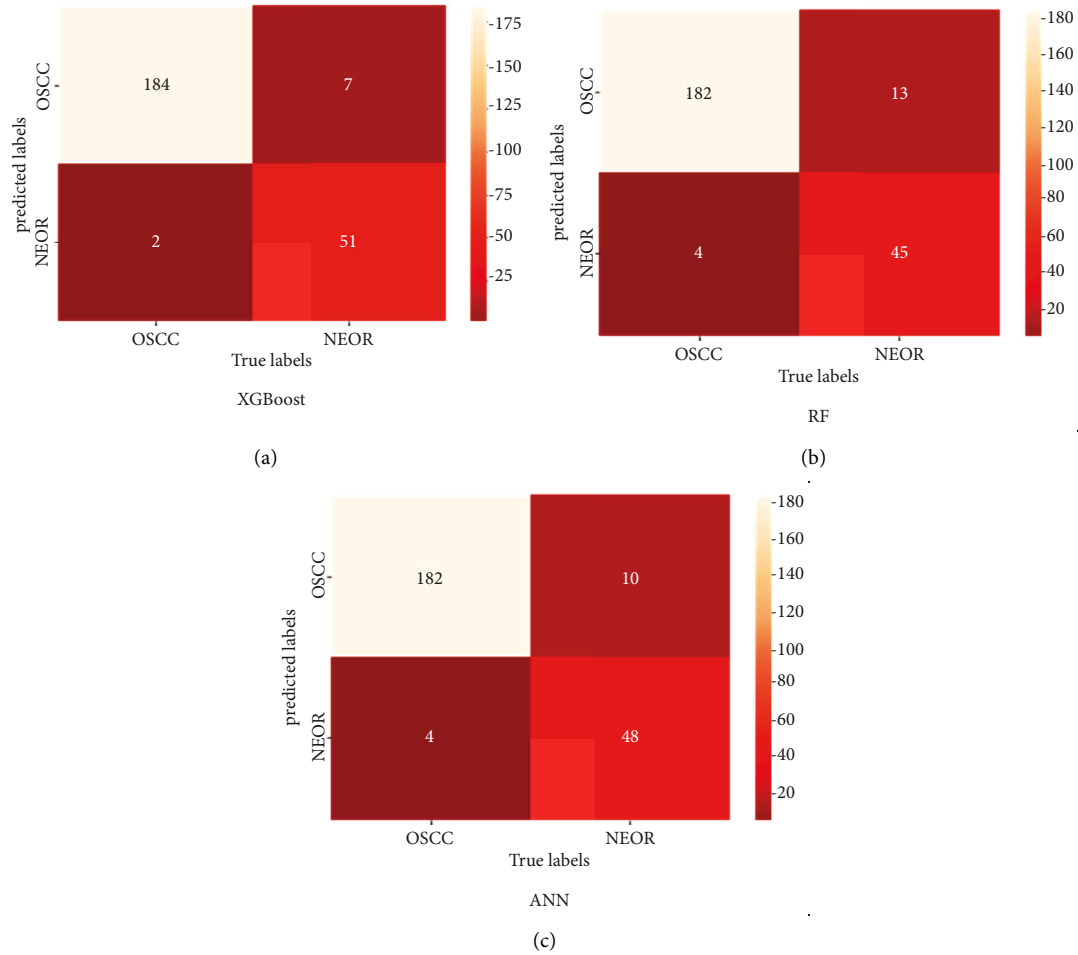
FIGURE 14: Confusion matrix of classification results for all classifiers.

TABLE 4: Accuracy performance for the classifier with non-stain-normalized image and stain-normalized image

Classifier name	Without stain normalization (%)	With stain normalization (%)
XGBoost	95.1	96.3
RF	91.3	93.1
ANN	94.2	94.1

186 histopathological images with OSCC were detected correctly, while two were diagnosed as NEOR. On other hand, the classifier had 7 histopathological images with NEOR. The XGBoost classifier has misclassified 7 normal histopathological images, which is the lowest number of FP compared to other classifiers.

According to the results shown in the confusion matrix, overall accuracy, sensitivity, and precision were calculated, as shown in Table 3. The XGBoost model delivered outstanding results compared to the other classifiers. It has consistently high accuracy, sensitivity, and precision (96.3%, 98.9%, and 96.3% respectively) across all models. This shows that this model was the most successful in learning and extracting essential features from the training data.

Finally, the proposed work investigates whether data normalization techniques can improve the accuracy of classification models or not, where all models were trained on nonnormalized histopathological images and the classification performance is illustrated in Table 4. It is seen that the accuracy was decreased when using histopathological images without applying the Reinhard approach. These effective results show that Reinhard stain normalization can improve the classifier performance and can produce satisfactory results.

For further evaluation of the proposed approach, it was compared with several previous research works shown in Table 5, which shows that the proposed method achieves the highest accuracy of 96.3% to classify different classes of histopathological images.

TABLE 5: Comparison of previous research works with the proposed method.

Authors	Algorithm architecture	The objective of the study	No. of images	Accuracy
Uthoff et al. [52]	CNN	Early detection of precancerous and cancerous lesions	170	86.88%
Jeyaraj et al. [53]	CNN	Develop an automated computer-aided oral cancer-detecting system	100	91.4%
Jubair et al. [54]	CNN	CNN using EfficientNet-B0 transfer model CNN for binary classication of oral lesions into benign and malignant	716	85.0%
Sunny et al. [55]	ANN	Early detection of oral potentially malignant/malignant lesion	82	86%
Rathod et al. [56]	CNN	Classify different stages of oral cancer	Not mentioned	90.68%
Alabi et al. [57]	ANNs	Predicting risk of recurrence of oral tongue squamous cell carcinoma (OTSCC)	311	81%
Alhazmi et al. [58]	ANNs	Predicting risk of developing oral cancer	73	78.95%
Chu et al. [59]	CNN	Evaluate the ability of supervised machine learning models to predict disease outcome	467	70.59%
Karadaghy et al. [60]	DT	Develop a prediction DT model using machine learning for 5-year overall survival among patients with OCSCC	33, 065	71%
Proposed method	CNN and BPSO	The features were extracted with Inception V3 and selected with BPSO to improve the classification performance of patients with OCSCC	1224	96.3%

6. Conclusion

In this study, traditional classification algorithms are used by extracting features from the four CNN models (VGG16, AlexNet, ResNet50, and Inception V3) and selecting the best features using the BPSO algorithm. The features extracted with Inception V3 and selected with BPSO improved the classification performance and contributed positively to the results. In addition, the effects of stain normalization procedures were investigated and compared with nonnormalized histological images. The results showed that the XGBoost model performed better when the Reinhard technique was used. This breakthrough achievement has the potential to be a precious and rapid diagnostic tool that might save many people who die each year because of delayed or incorrect diagnoses.

Data Availability

The data used to support the findings of the study can be obtained from the first author (e-mail: mohand.deif@eng.mti.edu.eg) upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest to declare.

Acknowledgments

References

[1] M. Du, R. Nair, L. Jamieson, Z. Liu, and P. Bi, "Incidence trends of lip, oral cavity, and pharyngeal cancers: global burden of disease 1990–2017," *Journal of Dental Research*, vol. 99, no. 2, pp. 143–151, 2020.

[2] L. Li, Y. Yin, F. Nan, and Z. Ma, "Circ_LPAR3 promotes the progression of oral squamous cell carcinoma (OSCC)," *Biochemical and Biophysical Research Communications*, vol. 589, pp. 215–222, 2022.

[3] S. Warnakulasuriya and J. S. Greenspan, "Epidemiology of oral and oropharyngeal cancers," in *Textbook of Oral Cancer*Springer, Berlin, Germany, 2020.

[4] S. Perdomo, G. Martin Roa, P. Brennan, D. Forman, and M. S. Sierra, "Head and neck cancer burden and preventive measures in central and South America," *Cancer Epidemiology*, vol. 44, pp. S43–S52, 2016.

[5] N. Anwar, S. Pervez, Q. Chundriger, S. Awan, T. Moatter, and T. S. Ali, "Oral cancer: clinicopathological features and associated risk factors in a high-risk population presenting to a major tertiary care center in Pakistan," *PloS One*, vol. 15, no. 8, Article ID e0236359, 2020.

[6] D. Chakraborty, C. Natarajan, and A. Mukherjee, "Advances in oral cancer detection," *Advances in Clinical Chemistry*, vol. 91, pp. 181–200, 2019.

[7] A. W. Eckert, M. Kappler, I. GroBe, C. Wickenhauser, and B. Seliger, "Current understanding of the HIF-1-Dependent metabolism in oral squamous cell carcinoma," *International Journal of Molecular Sciences*, vol. 21, no. 17, 2020.

[8] A. Ghosh, D. Chaudhuri, S. Adhikary, and A. K. Das, "Deep reinforced neural network model for cyto-spectroscopic analysis of epigenetic markers for automated oral cancer risk prediction," *Chemometrics and Intelligent Laboratory Systems*, vol. 224, Article ID 104548, 2022.

[9] M. A. Deif and R. E. Hammam, "Skin lesions classification based on deep learning approach," *Journal of Clinical Engineering*, vol. 45, no. 3, pp. 155–161, 2020.

[10] J. Kong, O. Sertel, H. Shimada, K. Boyer, J. Saltz, and M. Gurcan, "Computer-aided evaluation of neuroblastoma on whole-slide histology images: classifying grade of neuroblastic differentiation," *Pattern Recognition*, vol. 42, no. 6, pp. 1080–1092, 2009.

[11] M. F. Santana and L. C. L. Ferreira, "Diagnostic errors in surgical pathology," *Jornal Brasileiro de Patologia e Medicina Laboratorial*, vol. 53, pp. 124–129, 2017.

[12] M. A. Deif, R. E. Hammam, and A. A. A. Solyman, "Gradient boosting machine based on PSO for prediction of leukemia after a breast cancer diagnosis," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 11, no. 2, pp. 508–515, 2021.

[13] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: concepts, methods, challenges, and future directions," *IEEE Access*, vol. 7, Article ID 99540, 2019.

[14] M. A. Deif, A. A. A. Solyman, M. H. Alsharif, and P. Uthansakul, "Automated triage system for intensive care admissions during the COVID-19 pandemic using hybrid XGBoost-AHP approach," *Sensors*, vol. 21, no. 19, 2021.

[15] A. Echle, N. T. Rindtorff, T. J. Brinker, T. Luedde, A. T. Pearson, and J. N. Kather, "Deep learning in cancer pathology: a new generation of clinical biomarkers," *British Journal of Cancer*, vol. 124, no. 4, pp. 686–696, 2021.

[16] A. Duggento, A. Conti, A. Mauriello, M. Guerrisi, and N. Toschi, "Deep computational pathology in breast cancer," *Seminars in Cancer Biology*, vol. 72, pp. 226–237, 2021.

[17] S. Wang, D. M. Yang, R. Rong et al., "Artificial intelligence in lung cancer pathology image analysis," *Cancers*, vol. 11, no. 11, p. 1673, 2019.

[18] S. L. Goldenberg, G. Nir, and S. E. Salcudean, "A new era: artificial intelligence and machine learning in prostate cancer," *Nature Reviews Urology*, vol. 16, no. 7, pp. 391–403, 2019.

[19] D. K. Das, S. Bose, A. K. Maiti, B. Mitra, G. Mukherjee, and P. K. Dutta, "Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis," *Tissue and Cell*, vol. 53, pp. 111–119, 2018.

[20] N. Das, E. Hussain, and L. B. Mahanta, "Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network," *Neural Networks*, vol. 128, pp. 47–60, 2020.

[21] J. Folmsbee, X. Liu, M. Brandwein-Weber, and D. Scott, "Active deep learning: improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer," in *Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 770–773, Washington, DC, USA, April 2018.

[22] F. Martino, D. D. Bloisi, A. Pennisi et al., "Deep learning-based pixel-wise lesion segmentation on oral squamous cell carcinoma images," *Applied Sciences*, vol. 10, no. 22, 2020.

[23] I. Amin, H. Zamir, and F. F. Khan, "Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models," medRxiv, http://medrxiv.org/content/early/2021/05/14/2021.05.06.21256741. abstract, 2021.

[24] M. Abbasi, R. Tahouri, and M. Rafiee, "Enhancing the performance of the aggregated bit-vector algorithm in network packet classification using GPU," *PeerJ Computer Science*, vol. 5, 2019.

[25] M. Abbasi, S. Vesaghati Fazel, and M. Rafiee, "MBitCuts: optimal bit-level cutting in geometric space packet classification," *The Journal of Supercomputing*, vol. 76, no. 4, pp. 3105–3128, 2020.

[26] M. Abbasi and A. Shokrollahi, "Enhancing the performance of decision tree-based packet classification algorithms using

[27] CPU cluster," *Cluster Computing*, vol. 23, no. 4, pp. 3203–3219, 2020.

[27] T. Y. Rahman, "A histopathological image repository of normal epithelium of oral cavity and oral squamous cell carcinoma," 2019, https://onlinelibrary.wiley.com/doi/abs/10.1111/jmi.12611.

[28] M. A. Deif, "Deep learning algorithms and soft optimization methods codes," 2022, https://github.com/BioScince/Oral-squamous-cell-carcinoma-.git.

[29] M. A. Deif, A. A. A. Solyman, and R. E. Hammam, "ARIMA model estimation based on genetic algorithm for COVID-19 mortality rates," *International Journal of Information Technology & Decision Making*, vol. 20, no. 6, pp. 1775–1798, 2021.

[30] M. A. Deif, R. E. Hammam, A. A. A. Solyman, and S. S. Band, "A deep bidirectional recurrent neural network for identification of SARS-CoV-2 from viral genome sequences," *Mathematical Biosciences and Engineering*, AIMS--Press, vol. 18, no. 6, 2021.

[31] C. Munien and S. Viriri, "Classification of hematoxylin and eosin-stained breast cancer histology microscopy images using transfer learning with EfficientNets," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5580914, 17 pages, 2021.

[32] D. Magee, D. Treanor, D. Crellin, and M. Shires, "Colour normalisation in digital histopathology images," in *Proceedings of the Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy (MICCAI Workshop)*, pp. 100–111, 2009.

[33] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Applications*, vol. 21, no. 4, pp. 34–41, 2001.

[34] A. Janowczyk, A. Basavanhally, and A. Madabhushi, "Stain normalization using sparse autoencoders (StaNoSA): application to digital pathology," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 50–61, 2017.

[35] M. Tan and Q. Le, "Efficientnet: rethinking model scaling for convolutional neural networks," *International Conference on Machine Learning*, pp. 6105–6114, 2019.

[36] M. Macenko, M. Niethammer, J. S. Marron, and D. Borland, "A method for normalizing histology slides for quantitative analysis," in *Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Boston, MA, USA, June 2009.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[40] C. Szegedy, V. Vanhoucke, and S. Loffe, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[41] M. Canayaz, "MH-COVIDNet: diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images," *Biomedical Signal Processing and Control*, vol. 64, Article ID 102257, 2021.

[42] A. Narin, "Accurate detection of COVID-19 using deep features based on X-ray images and feature selection

methods," *Computers in Biology and Medicine*, vol. 137, Article ID 104771, 2021.

[43] M. A. Deif, H. Attar, A. Amer, H. Issa, M. R. Khosravi, and A. A. A. Solyman, "A new feature selection method based on hybrid approach for colorectal cancer histology classification," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7614264, 14 pages, 2022.

[44] M. A. Deif, A. A. A. Solyman, M. H. Alsharif, S. Jung, and E. Hwang, "A hybrid multi-objective optimizer-based svm model for enhancing numerical weather prediction: a study for the seoul metropolitan area," *Sustainability*, vol. 14, no. 1, 2021.

[45] K. Kumari, J. P. Singh, Y. K. Dwivedi, and N. P. Rana, "Multimodal aggression identification using convolutional neural network and binary particle swarm optimization," *Future Generation Computer Systems*, vol. 118, pp. 187–197, 2021.

[46] R. E. Hammam, H. Attar, A. Amer, and H. Issa, "Prediction of wear rates of UHMWPE bearing in hip joint prosthesis with support vector model and grey wolf optimization," *Application of Neural Network in Mobile Edge Computing*, vol. 2022, Article ID 6548800, 16 pages, 2022.

[47] M. Deif, R. Hammam, and S. Ahmed, "Adaptive neuro-fuzzy inference system (ANFIS) for rapid diagnosis of COVID-19 cases based on routine blood tests," *International Journal of Intelligent Engineering and Systems*, vol. 14, 2021.

[48] F. R. V. Alves and R. P. Machado Vieira, "The Newton fractal's leonardo sequence study with the Google Colab," *International Electronic Journal of Mathematics Education*, vol. 15, no. 2, Article ID em0575, 2019.

[49] R. G. Van, "Python programming language," in *Proceedings of the USENIX Annual Technical Conference*, 2007.

[50] A. Gulli and S. Pal, *Deep Learning with Keras*, Packt Publishing Ltd, Birmingham, United Kingdom, 2017.

[51] S. B. Khanagar, S. Naik, A. A. Al Kheraif et al., "Application and performance of artificial intelligence technology in oral cancer diagnosis and prediction of prognosis: a systematic review," *Diagnostics*, vol. 11, no. 6, 2021.

[52] R. D. Uthoff, B. Song, S. Sunny et al., "Point-of-Care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities," *PloS One*, vol. 13, no. 12, Article ID e0207493, 2018.

[53] P. R. Jeyaraj, E. R. Samuel Nadar, and R. S. N. Edward, "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm," *Journal of Cancer Research and Clinical Oncology*, vol. 145, no. 4, pp. 829–837, 2019.

[54] F. Jubair, O. Al-karadsheh, D. Malamos, S. Al Mahdi, Y. Saad, and Y. Hassona, "A novel lightweight deep convolutional neural network for early detection of oral cancer," *Oral Diseases*, vol. 28, no. 4, pp. 1123–1130, 2022.

[55] S. Sunny, A. Baby, B. L. James et al., "A smart tele-cytology point-of-care platform for oral cancer screening," *PLoS One*, vol. 14, no. 11, Article ID e0224885, 2019.

[56] J. Rathod, S. Sherkay, H. Bondre, and R. Sonewane, "Oral cancer detection and level classification through machine learning," *Int. J. Adv. Res. Comput. Commun. Eng*, vol. 9, pp. 177–182, 2020.

[57] R. O. Alabi, M. Elmusrati, I. Sawazaki-Calone et al., "Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer," *International Journal of Medical Informatics*, vol. 136, Article ID 104068, 2020.

[58] A. Alhazmi, Y. Alhazmi, A. Makrami et al., "Application of artificial intelligence and machine learning for prediction of oral cancer risk," *Journal of Oral Pathology & Medicine*, vol. 50, no. 5, pp. 444–450, 2021.

[59] C. S. Chu, N. P. Lee, J. Adeoye, P. Thomson, and S. Choi, "Machine learning and treatment outcome prediction for oral cancer," *Journal of Oral Pathology & Medicine*, vol. 49, no. 10, pp. 977–985, 2020.

[60] O. A. Karadaghy, M. Shew, J. New, and A. M. Bur, "Development and assessment of a machine learning model to help predict survival among patients with oral squamous cell carcinoma," *JAMA Otolaryngology–Head & Neck Surgery*, vol. 145, no. 12, pp. 1115–1120, 2019.

Hindawi

Research Article

Optimization Design of Short Life Cycle Product Logistics Supply Chain Scheme Based on Support Vector Machine

Foshang Li [ID]

School of Business, Jilin Business and Technology College, Jilin, Changchun 130507, China

Correspondence should be addressed to Foshang Li; 2015223010033@stu.scu.edu.cn

In order to better solve the problem of product logistics supply chains with short life cycles, a solution optimization of short life cycle product logistics supply chains based on support vector machines is proposed. This method recommends key technical problems and solutions through information represented by support vector machines and explore the research of short life cycle products to realize logistics supply chains. The research shows that, whether it is a retail channel or a network channel, the RMSE value of the effect index predicted by SVM is smaller than the RMSE value of the improved Bass. It can be seen that the SVM demand forecasting model constructed by considering multiple input factors can obtain a more accurate forecasting effect. The accuracy of the demand forecasting model based on SVM is verified.

1. Introduction

In recent years, the rapid development of the e-commerce industry not only drives the rapid development of the traditional consumption field but also gradually forms the dual-channel supply chain mode of coexistence of the traditional industry and the e-commerce platform. The rapid development of e-commerce directly promotes the rapid development of the traditional logistics industry. In addition, the continuous progress of science and technology, the transformation and rise of personalized demand of market consumer groups have virtually promoted the upgrading of products, and at the same time, the life cycle of products is constantly shortened; that is, short life cycle products have begun to appear. The market demand of this kind for product is too uncertain and time-sensitive, which virtually increases the difficulty of the supply chain management of enterprises. Enterprise supply chain management requires the mutual activities of different entities and related functions to gradually achieve a high degree of coordination, and on the basis of a high degree of coordination and cooperation, gradually achieve the purpose of a more flexible and low-cost corresponding consumer market [1, 2]. Therefore, the product supply chain with a short life cycle must be

further optimized. On this basis, a model prediction algorithm based on support vector machine and random inventory control adjustment and an optimization strategy based on TOC dynamic buffer management advantage are proposed, and this model is applied to the short life cycle product logistics supply chain to optimize the supply chain scheme [3]. Due to the fashion, seasonality, and depreciation characteristics of short life cycle products, it is particularly important to accurately predict logistics supply chain information and quickly respond to market changes. Traditional time series and other forecasting methods need to find out the law of historical data changing with time to predict market demand, but this method lacks consideration of short product life cycles and lacks historical data. Some scholars take into account the characteristics of short life cycle products and use the Bass model family to achieve prediction, but there is a lack of research on the different characteristics of traditional and network channels. The support vector machine (SVM) method has significant advantages in solving the problem of small sample and nonlinear high-latitude model construction, and can improve the generalization ability of the model. Therefore, the use of SVM for forecasting has lower requirements on the amount of data, and is suitable for demand forecasting of

products with short life cycles. Combining the characteristics of short life cycle product supply chain dual-channel sales cycle and less historical sales data, this paper discusses a short life cycle product demand forecast model based on Bass model that only considers the single factor of sales historical data; analyzes service and purchase intentions Bass improvement model under two factors. Based on this, the SVM algorithm is applied to the demand forecast of short life cycle products under dual-channel, considering the influence of price, season, promotion, industry status, historical sales data, and channel type of the short life cycle products in dual-channel sales. Combining the forecast correction value of the Bass model family with various influencing factors such as channels, a demand forecast model based on SVM is constructed, and simulation analysis shows that this method can improve the effect of short life cycle product dual-channel supply chain demand forecasting.

2. Related Works

Due to the rapid change in the market environment and various internal problems in the supply chain, the supply chain system must deal with various possible uncertainties in supply and demand in an appropriate way. At the same time, due to the inventory in the supply chain system, between the link between the medium and the buffer role. Therefore, good supply chain inventory management can effectively reduce the impact of various uncertainties on the overall supply chain system [4]. At the same time, in supply chain inventory management, a better management strategy and method should be able to ensure a higher level of customer service it can also maintain reasonable and effective control over the overall inventory level of the supply chain system, and increase the overall effective output of the system as much as possible, and further improve the comprehensive competitiveness of enterprises in the supply chain [5, 6]. Therefore, inventory management has always been one of the important issues in traditional supply chain research, and related research has been continuously developing. The existing inventory theory research can be roughly divided into deterministic inventory models, stochastic inventory models, multilevel inventory model, infinite cycle inventory models, and various inventory contract models which are currently in the research hotspot. For the deterministic inventory model, its main parameters, including the specific order quantity of each batch, order cost, and unit inventory cost, remain unchanged while the demand for related products is determined. In the research of this model framework, we generally conduct in-depth analysis and research from the perspectives of product quantity discount setting, replenishment lead time setting, and how to compensate for out-of-stock. In reality, there are many conditions, such as uncertain lead time, random demand, and variable ordering costs, and out-of-stock allowance. Therefore, the research on the stochastic inventory models has more practical significance [7].

Short life cycle products are commodities with a relatively short and fixed sales time, such as newspapers and magazines, fashion, all kinds of decorations, and increasingly electronic consumer goods [8]. For these short life products, distributors typically place orders well in advance of the selling season. It is also difficult for distributors to obtain particularly accurate demand information before placing orders. Therefore, only after calculating the expected revenue brought by all kinds of possible situations, such as shortage or overstocking, can it formulate the optimal ordering strategy by way of prediction according to the actual situation.

Considering that manufacturers' channel selection, channel competition, and contract coordination are the main factors affecting supply chain performance, some scholars put forward that perishable manufacturers' sales channel selection under the Internet environment will be affected by product price, the freshness of perishable products, and product price and quality of competitive channels. It is found that the increase of e-commerce channels is beneficial to the profit increase of the whole supply chain system. In the case of high distribution cost, the low wholesale price, and market demand fluctuation, it is more advantageous to choose two sales channels than a single channel. Considering the convenience factor of the channel, the pricing problem of the two sales channels is studied after the manufacturer adds the e-commerce channel to the two-layer supply chain system. The optimal coordinated price strategy of a dual-source supply chain is studied from the perspective of whether demand information is shared. The influence of service competition on channel pricing and profit distribution in the dual-channel supply chain is studied. From the perspective of consumers' sensitivity to service quality and price, different coordination modes of supply chain dual-source channels are studied. This article studies the win-win problem of supply chain channels achieved by manufacturers through quantity discount contracts in a mixed channel supply chain systems [9, 10].

3. Methods

3.1. Support Vector Machine Classification Algorithm. In the case of linear separability, a support vector machine is proposed based on the optimal classification hyperplane of a binary classification problem [11]. For example, if we look at two separate classes of problems on a two-dimensional plane, the basic idea can be expressed as a two-dimensional planar graph (see Figure 1). It is easy to observe that the two categories are divided into many possible boundaries, and each boundary has a related edge. If we choose the one with the maximum interval, there is little possibility of mistakes in the classification of unknown items in the future. Support vector machines map samples from input space to higher dimensional space by the kernel function. In this way, the samples on both sides of the dividing boundary can be accurately split, and the optimal classification plane can be found to maximize the classification distance between the two categories [12]. The empirical risk $\text{Remp}(w) = 0$ and the confidence range $\varphi(h/n)$ is minimized, so that the real risk $R(w)$ is minimized.
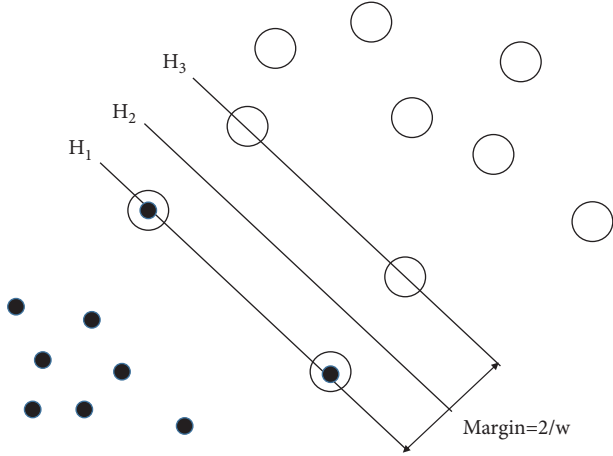
FIGURE 1: Optimal classification surface.

Assume that the given sample set is shown in:

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\},$$
$$x \in R^d, \quad (1)$$
$$y \in \{1, -1\}.$$

The two types of samples in the figure above are represented by hollow points and solid points, respectively. $H2$ is the classification line, i.e. there is no boundary line separating the two types incorrectly, and the width of $H2$ and $H3$ lines is the classification interval [13]. 。 The normalized processing, equation $w \cdot x + b = 0$ for the equation of classification line, in high dimensional feature space, can be calculated by the equation $H2$, $H3$ classification interval is $2/\|w\|$. According to the theory in the previous section, in order to obtain the maximum interval hyperplane of a linearly separable problem, constraints should be satisfied as shown in equation:

$$y_i (w \cdot x_i + b) \geq 1. \quad (2)$$

To maximize the interval $\triangle$ under the precondition of satisfying the above conditions, it is equivalent to the following optimization problem as shown in equations (3) and (4):

$$\min_{wh} \frac{1}{2} \|w\| s.t, \quad (3)$$

$$y_i (w \cdot x_i + b) \geq 1, i = 1, \ldots, l. \quad (4)$$

The above equation is a typical linear constrained convex quadratic programming problem, which determines the unique maximum interval classification hyperplane. Its Lagrange function is shown in the following equation:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - l \sum \alpha_i (y_i (w \cdot x_i + b) - 1). \quad (5)$$

The variable α_i is the Lagrange multiplier. The saddle-point control optimization problem of the Lagrange function $L(w, b, \alpha)$, which differentiates with respect to b and w. $L(w, b, \alpha)$ solved the differential equations of w and b respectively and made the results zero. Two optimization conditions could be obtained, as shown in equations (6) and (7):

$$\text{Condition1:} \frac{\partial L(w, b, \alpha)}{\partial w} = 0, \quad (6)$$

$$\text{Condition2:} \frac{\partial L(w, b, \alpha)}{\partial b} = 0. \quad (7)$$

Equation (8) can be obtained from condition 1:

$$w = \sum_{i=1}^{l} \alpha_i y_i x_i. \quad (8)$$

That is, the linear combination of training sample vectors is the weight vector of the optimal hyperplane. According to condition 2, the following equation can be obtained:

$$\sum_{i=1}^{l} \alpha_i y_i = 0. \quad (9)$$

Equation (8) problem can be transformed into a simpler "dual" problem by equation (9), as shown in the following equation:

$$W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i, x_j). \quad (10)$$

Equation (10) is the extremum problem of quadratic functions with inequality constraints. According to the optimality condition—Karush-Kühn-Tucke condition, the solution of the optimization problem of equation (10) must satisfy:

$$\alpha_i \{[(w \cdot x_i) + b] y_i - 1\} = 0, i = 1, \ldots, l. \quad (11)$$

After solving the above quadratic programming problem, the classified decision function can be expressed asfollows:

$$f(x) = \text{sgn}\left(n \sum \alpha_i * y_i (x_i \cdot x) + b^*\right). \quad (12)$$

When the training sample set is linear and not time-sharing, nonnegative relaxation variables can be introduced, as shown in:

$$\xi_i, i = 1, 2, \ldots, l. \quad (13)$$

The optimal problem of classified hyperplane is shown in:

$$\min_{w, b, \xi_i} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{1} \xi_i. \quad (14)$$

Solving the maximum value of the following function for α is its duality problem, as shown in equations (15)– and (17):

$$\sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i, x_j), \quad (15)$$

$$y_i (w \cdot x_i + b) \geq 1 - \xi_i, i = 1, \ldots, l, \quad (16)$$
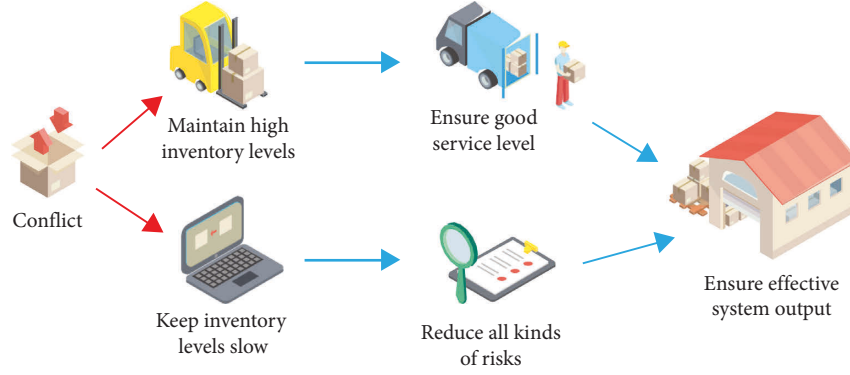
Figure 2: Inventory control system conflict diagram.

$$\sum_{i=1}^{l} y_i \alpha_i = 0,$$
$$0 \le \alpha_i \le C, \qquad (17)$$
$$\xi_i \ge 0, \quad i = 1, \ldots, l.$$

The performance of SVM is largely determined by the choice of penalty parameters and kernel parameters. In recent years, many scholars have focused on how to determine the parameters of kernel function to obtain the optimal generalization ability of SVM. At present, the simplest method is to first set the interval, range of each parameter, and search exhaustively within the defined step interval, and test one by one to get the parameter with the best performance on the learning object. In the application of the SVM model, the determination of the SVM model is determined by the determination of the kernel function. Therefore, the selection of kernel parameters is a very key step in the establishment of the SVM model. Although there is no universally accepted the best guideline for kernel selection, the Gaussian radial basis (RBF) kernel is the most commonly used kernel in many experiments and studies. Similarly, the radial basis kernel function (RBF) was selected as the kernel function of SVM. By analyzing the influence of radial basis parameters on the regression performance of the SVM model, the decision basis for the optimization of the SVM algorithm can be provided [14, 15].

3.2. Inventory Buffer Control Method in Supply Chain Coordination

3.2.1. Questions Raised. With the increasingly fierce market competition, customer preferences for short life cycle product supply chains are constantly changing. The increasing variety of commodities, the shorter life cycle of products, the frequent fluctuation of demand, and many other factors make the demand of each link of the whole supply chain show strong uncertainty and nonstability, and the uncertainty of demand greatly increases the difficulty of inventory control and management. According to the characteristics of short life cycle product supply chain, TOC puts forward methods and suggestions to establish a corresponding coordination system. The first is to take the VMI

inventory system as the "link" between all the links in the supply chain, and apply the buffer management strategy in TOC, and then take this as the basis and main way of information transmission between each link. Therefore, in fact, in this supply chain system, the inventory system determines the overall "strength" of the supply chain system, so the setting of the inventory control system is particularly important for the short life cycle product supply chain based on the TOC system to be studied [16, 17]. The main contradictions commonly existing in various supply chain inventory control systems are shown in Figure 2.

3.2.2. Problem Analysis. On the one hand, in order to ensure a good service level for the downstream links and reduce possible delivery delays or shortages, it is necessary to maintain a high inventory level of raw material parts or finished products in the supply chain system. On the other hand, upstream supply chain enterprises also need to reduce the backlog of raw material parts inventory or finished products, the occupation of capital and all kinds of risks brought by inventory. Therefore, it is necessary to keep a small inventory [18]. At the same time, it should be noted that in the case of nonstationary demand, the contradiction between inventory cost and production satisfaction rate is often more difficult to balance because the fluctuation of demand is difficult to be described by probability model. In order to ensure the effective output of the whole supply chain system, it is necessary to ensure the better service level of each link of the system and reduce the various risks faced by the supply chain system. The assumption of the upper part of the conflict diagram is that if the overall inventory of the supply chain keeps a high inventory level, good service level of the supply chain system can be guaranteed [19]. However, the current reality is that the coexistence of high inventory level and a low service level often leads to the shortage of urgently needed products in the market and the high inventory level of products with poor sales. Therefore, a higher inventory level cannot guarantee a good service level in the supply chain system. The assumptions therein are not tenable similarly, for the lower half of the conflict graph; it is assumed that if the overall inventory of the supply chain keeps a low inventory level, various risks caused by inventory can be reduced in the supply chain system [20].

Especially in the supply chain of fashion products, the demand volatility is often very large, especially in some seasonal peak demand; too low inventory level will make the supply chain system unable to meet the market demand, and then, have to face the risk of more severe market competition. Therefore, a low inventory level cannot completely reduce various risks in the supply chain system, and its hypothesis is also not valid. Based on the above analysis, in fact, the core problem of inventory control is still caused by the uncertainty of demand. If the inventory level is maintained at the appropriate time, the core conflict will be resolved, so the supply chain coordination needs a set of inventory control systems that can dynamically adjust the inventory level.

3.2.3. Problem Solving. In order to carry out effective inventory control in supply chain coordination and solve the contradiction between inventory cost and customer service level, the following measures can be taken: first, improve the accuracy of demand prediction, make full use of historical demand data and expert knowledge provided by decision makers to make an effective prediction. Second, the inventory control strategy should have strong flexibility. When demand changes suddenly or seasonal demand comes, inventory control parameters can be flexibly adjusted according to historical information.

3.3. Construction of an Improved Bass Demand Forecast Model Based on Sales Records. The demand forecasting method, usually based on historical sales records, is to scientifically fit the past sales through a large number of historical sales records of products to find their demand trends. Due to the short sales period of short life cycle products, it is difficult to obtain a large number of historical sales records for a long period of time, usually only one year or even a few months of sales records. In response to this situation, this paper regards market sales as a process of continuous diffusion and penetration. That is, first, dealers attract a small number of consumers with a certain sense of innovation through product promotion or trial, and these consumers drive subsequent consumers, and so on. While a realistic diffusion process involves a large number of factors and their relationships, this complexity can be modeled and studied by the improved Bass diffusion model. Considering that the improved Bass model family has better applicability, the improved Bass model is firstly used to discuss the demand forecasting problem of short life cycle products, considering only the historical sales record as a single factor. The Improved Bass model is a model used to predict the sales of consumer goods. The improved Bass model can analyze the decision time of consumers to purchase products. It believes that the buyers of a product are influenced by external or internal factors, and thus, divides potential buyers of products into two categories: (1) innovation groups, which are susceptible to external influences, that is, mass media; (2) imitation groups, which class groups are susceptible to internal influence, that is, word-of-mouth influence. The core idea of the model is that the purchasing decision of the
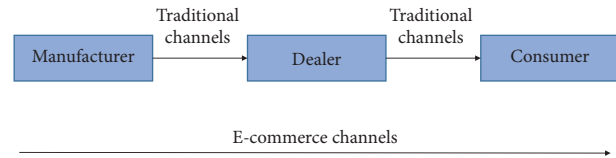
Figure 3: E-commerce dual-channel supply chain system.

innovation group is independent of the purchasing behavior of other members of the social system, and the time when the imitation group purchases products is affected by the social system, and this influence increases with the increase of the number of buyers, thus, imitating the purchasing decision of the group. Time is influenced by members of the social system. According to the Bass thought above, let $f(t)$ be the time density function of adopters in period t, indicating the possibility of purchasing in period t; $f(t)$ is the proportion of cumulative adopters to all buyers in period t; ρ is the innovation group coefficient, indicating the external influence of the attributes of the communication product that can be easily verified, such as price, function, and media $\rho \in (0, 1)$; The propaganda function of potential users, such as products, disseminates some characteristics of products that require long-term experience to discover, $q \in (0, 1)$; m is the potential purchase amount; $n(t)$ is the cumulative purchase amount in the t period; $n(t)$ is the noncumulative purchase amount in period t; $X(i)$ is the product demand.

3.4. Optimization Model of Short Life Cycle Product Logistics Supply Chain Scheme Based on Support Vector Machine. The demand forecasting method based on historical sales records only focuses on the sales records of products, so it has certain limitations. The demand forecasting method based on service-purchase intention takes into account the market behavior characteristics of participants in addition to the sales record data, and has made certain progress. However, these two schemes can only distinguish the two channels by their different historical sales record values, but ignore the unique characteristics of the two channels. To this end, this paper considers the dual-channel demand forecast for short life cycle products under multiple influencing factors. Multiinfluencing factors refer to the dual-channel market demand that is affected by multiple factors. For retail channels, it is mainly affected by product price, seasonal index, promotion level index, industry status quo, the number of people entering the store, and historical sales data; for online channels, it mainly considers product price, seasonal index, and promotion level index, industry status, credit index and historical sales data and other influencing factors. Considering that SVM can effectively analyze multiple small samples and multiple input factors, the SVM method is combined with the improved Bass diffusion model, and the product price, seasonal index, promotion degree index, industry status quo, number of shoppers, credit index, and historical sales are introduced. Data and other market-influencing factors are used, and the sales value predicted by the improved Bass model is used as the

input vector to establish a dual-channel short life cycle product demand forecasting model under multiple influencing factors.

3.4.1. Establish a Demand Forecasting Model Based on Support Vector Machine.

According to the basic idea of SVM regression, through a nonlinear mapping Φ, the sample data (x_i, y_i) is mapped to the high-latitude feature space F, and linear regression is performed in this space:

$$f(x) = \left(\alpha^T \Phi(x)\right) + b, \Phi: R^n \longrightarrow F, \alpha \in F. \quad (18)$$

The support vector machine regression can be expressed as the following constrained optimization problem, that is, the quadratic programming form of the objective function is

$$s, t. R^T \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix} = 0, 0 \le \alpha_i, \alpha_i \le C. \quad (19)$$

In the formula: $i = 1, 2, \ldots, m$; α is the weight vector, $\alpha \in F$; Q and P are the two specified vector groups. When $i = 1, 2, \ldots, l$, $r_i = 1$; then, $i = l + 1, l + 2, \ldots, 2l$, $r_i = -1$. Among them, $x_i (i = 1, 2, \cdots, l)$ is the input of the ith training sample, $y_i (i = 1, 2, \ldots, l)$ is the output of the ith training sample, write a program according to the above process, that is, get the regression decision function of the SVM demand forecasting model:

$$f(x) = \sum_{i=1}^{l} \left(\overline{\alpha}_i, \overline{\alpha}_i^*\right) K(x, x_i) + \overline{b}. \quad (20)$$

3.4.2. Determining SVM Optimization Parameters.

The most commonly used method for SVM parameter optimization is to make the minimum penalty parameter c and the reciprocal g of the number of attributes in the input data that can achieve the highest verification classification accuracy take values within a certain range. For a certain set of c and g, the K-CV method is used to obtain the training set verification classification accuracy under the set of c and g, and finally, the set of c and g with the highest training set verification classification accuracy is selected as the best. Parameters. After the regression decision function is obtained, the normalized forecast sample is substituted into the regression decision equation, that is, the value y of the demand forecast between $(0, 1)$ can be obtained, which can be converted into the actual forecast value by using the formula (20):

$$y = \widehat{y}(y_{\max} - y_{\min}) + y_{\min}. \quad (21)$$

In the above forecasting process, the fitted values, historical record values, and other market characteristics of the improved Bass model are substituted into the model in the form of input vectors, and the optimal parameters and demand forecast values can be determined through model simulation.

3.5. Dual-Channel Supply Chain Coupling Model for Short Life Cycle Products

3.5.1. Supply Chain Model of Different Economic Forms

(1) Supply Chain Model of Different Economic Forms Based on Customer Order Separation Point (CODP). For manufacturers, generally speaking, their internal supply chain operation is separated from customer orders, and there is a combination of design process, manufacturing process, assembly process, delivery process, and after-sales service process in the customer order completion process. Among them, the activity on the left of the customer order separation point is push type, and the activity on the right of the customer order separation point is a pull type. Therefore, it divides the supply chain into two parts: one is to respond directly to the needs of the end customer, and the other is to cushion the change of supply chain demand with planning and strategic inventory. The decoupling point is also the turning point of enterprise production activities from forecasted inventory production to customized production in response to customer demand. The equation and optimization of plans at this point is no longer based on prediction but on customer orders and resource allocation of the enterprise itself.

(a) *Engineer to Order (ETO).* The needs of each consumer are different. Production can only be organized by designing products according to the needs of consumers when they receive orders, which requires a long time to prepare production technology. In this type of production, products are largely designed to meet the requirements of a particular consumer. To be designed to support consumerization is an important function and component of the production process. In this type of production, the production batches of products are small, but the design work and the final product are often very complex. In the production process, each job has to be specially handled so that the various subparts of a large product or project can be precisely matched at the final stage so that the different subparts produced by different people and different places can be combined into a complex product or project. Very advanced configuration systems are required for overall coordination and management control.

(b) *Make to Order (MTO).* It purchases raw materials in advance according to product demand characteristics and forecast, and starts production after consumer demand is determined. The enterprise starts the manufacturing process after the order arrives, and the customer provides the design data, and the enterprise carries out personalized production according to the design requirements. Because customer orders trigger production, the finished product is rarely kept.

(c) *Assemble to Order (ATO).* A product is a sequence of optional parts or prefinished components that users can select according to their individual needs.

TABLE 1: Operation characteristics of lean supply chain and agile supply chain.

	Lean supply chain	Agile supply chain
Product design strategy	Achieve maximum performance with the lowest product cost	Provide solutions for customers, use a large number of common parts, and delay customer differentiation
The pricing strategy	The margin is low and price is the main factor to attract customers	Price is not the main attraction for customers
Production strategy	Reduce costs by increasing the utilization of equipment and personnel	Maintain production flexibility and try to meet uncertain demand
Inventory strategy	Realize the minimum inventory to array low-cost	Maintain safety inventory to meet uncertain demand
Lead-time strategy	On the basis of not increasing the cost, reduce the supply cycle	Maintain safety stock. To satisfy the uncertainty of demand
Supply strategy	Balance costs and service levels	Water in speed, elasticity, and service. Balance up
Transport strategy	More reliance on low-cost means of transportation	Rely more on block agility

Manufacturers produce semifinished products or components, assemble them according to customer orders, and reduce lead times by producing standardized modules and assembling customized products to order. It also reduces the risk of early production. For example, computer production and assembly manufacturers cannot assemble a computer without receiving an order. However, in order to improve the response speed to consumers, they can produce some standard components in advance according to the market forecast and conduct rapid assembly when the orders arrive to meet the personalized needs of consumers.

(d) *Make to Stock (MTS)*. Consumers basically have no suggestions or requirements for the determination of final product specifications, and they have little input. Manufacturers do not make products tailored to any particular consumer. The products of the enterprise are manufactured before the order arrives, and the customer needs are satisfied through the inventory. The MTS production plan is based on a market forecast; such products should have relatively accurate forecast accuracy.

(e) *Physical Distribution on Stock (PDS)*. The demand for production comes from distributors. The product is standardized, and the manufacturer does not know who the final consumer is during the production process but predicts demand information based on historical sales data. This is generally a mass-consumer product, and management focuses on forecasting and inventory control.

(2) Lean-Agile Supply Chain Coupling Based on Customer Order Separation Point (CODP). CODP is a certain point in the manufacturing process as well as a certain point in the supply chain, which is closely related to the enterprise supply chain. If the customer order separation point (CODP) of each link in a manufacturing enterprise is extended to a multistage supply chain, it involves suppliers, manufacturers, distributors, consumers, and other links. And for a supply chain system, its response time refers to the sum of
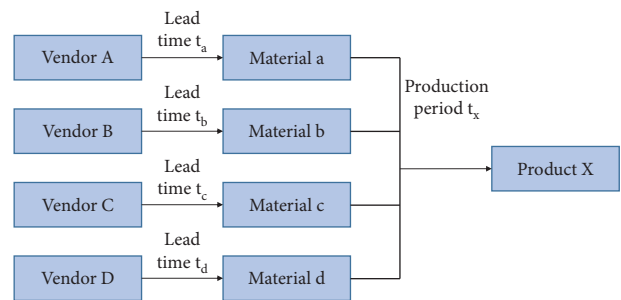
FIGURE 4: MRP material procurement model.

time accumulated in all stages of the whole process, from the supply chain detecting the demand of the end customer and receiving the order to the delivery of the product/service to the end customer. Obviously, the response process of the supply chain to the final customer is actually the completion process, of the customer order, and its length depends on the point at which the customer order customization begins. In general, the part of direct response to consumer demand has strong demand correlation and autocorrelation, and it is easy to complete its relatively accurate prediction with a scientific algorithm through traditional forecasting methods and adopt the supply chain operation mode of MTS. The other part needs to use demand-side supply chain predictive management and adopt the supply chain operation mode of MTO. The dual-channel supply chain system of e-commerce is shown in Figure 3.

3.5.2. The Delay Strategy is Coupled to the MTO/MTS Supply Chain

(1) MTO/MTS supply chain coupling model from the perspective of flexible low-cost response. Lean supply chain management originates from lean production management and its bigger foothold lies in the reduction of supply chain costs. It refers to a series of supply chain planning, implementation, and control processes that optimize and transform the whole chain, including the upstream and downstream chains, to eliminate unnecessary steps, delays, waiting, and consumption, eliminate waste in
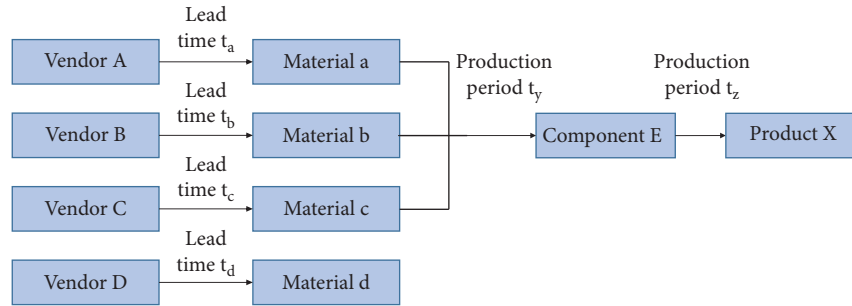
FIGURE 5: Delayed manufacturing model.

TABLE 3: Comparison of forecast—driven and order—pulled production economics.

Project	Predict to promote	Order pull
Starting point	Reduce purchasing and production cost as much as possible with accurate forecast	Be flexible and responsive to customer needs and changes
Location	Vendor-coupling point	Coupling point—consumer
Material characteristics	Long procurement cycle, variety, and centralized procurement to reduce unit price and production into components	Species is less
Processing requirements	Specialized equipment and skilled workers are generally required	Low requirements for equipment and workers
Supplier selection	Cost and quality	Speed, elasticity, and mass
Implementation method	Standardization, process, and modular	Customized
Economic benefits	Low cost operation	The pursuit of fast response can lead to higher costs
Advantages	Material versatility	Targeted, to meet the needs of customization
Disadvantages	Response period length	Need an inventory to be a buffer

enterprises, minimize costs, and meet consumer needs to the greatest extent. Lean supply chain management requires upstream and downstream efforts to reduce costs and waste throughout the process. See Table 1.

The competition of consumer demand response has developed from the simple quick delivery to the stage where enterprises unconditionally accept the increase, decrease, change or even cancel the demand put forward by consumers at any time. Enterprise cost management has also developed from the reduction of procurement, production, storage, transportation, and other physical costs to the total cost competition stage, including physical costs and market response costs, including out-of-sale costs and inventory holding costs. The above new supply chain competition stage generates a new supply chain development model—a flexible low-cost response model.

Combined with the above analysis and flexible low-cost response model, there is a coupling mechanism between lean supply chain and agile supply chain in enterprise operation. The advantages, disadvantages and complementarities of MTS-oriented supply chain strategy design in lean supply chain and MTO-oriented supply chain strategy design in agile supply chain are considered, and the factors and requirements of MTS supply chain in cost reduction and MTO supply chain in flexible response are combined.

(2) MTO/MTS supply chain coupling model based on two-stage supply system. Generally speaking, manufacturing enterprises issue material purchase orders based on MRP operations and require all materials to be delivered at the same time specified. Material quantity conforms to BOM structure ratio of production product. In order to facilitate the study, this section selects a simple material procurement model consisting of only four raw materials and only one supplier for each raw material, is shown in Figure 4.

In operation, the materials required for production are often not all in accordance with BOM requirements. When the delivery time of products is tight, the manufacturer will start production activities based on all the materials in stock when the delivery time of the missing materials is relatively clear. This is not only the practice of advance production but also the concrete implementation of delayed production, as shown in Figure 5.

Through the second-order supply system, manufacturers can divide a wide variety of raw materials with different lead periods into multiple components or materials (raw materials) with strong generality, relatively easy demand

TABLE 2: Supply chain based on time dimension Decoupling Points.

Supply chain structure	Flow of the supply chain	Process involved in each link
Manufacturers	Product design	(1) Modular design of components (2) BOM design
	Purchasing supply	(1) Material procurement process in different cycles (2) Different material procurement and inventory strategies
	Order processing	(1) Order acceptance and review (2) Production planning and control
	Production and processing	(1) Production process route
Dealers	Shipment delivery	(1) Order priority
	Order processing	(1) Order acceptance and review
	Shipment delivery	(1) Order priority
Consumers	Demand forecasting	(1) Prediction based on historical requirements (2) Requirement change
	Purchase order	(1) Regular or quantitative ordering plan
	Warehouse inventory	(1) Daily inventory strategy and level (2) Product acceptance and storage management

prediction, and a long production cycle according to the process characteristics, and it adopts an MTS type supply chain to compress the order response speed of this part of component modules or materials, and it purchases and manufactures costs to start final product assembly after receiving consumer orders, which not only meets the rapid demand of consumers but also reduces the cost of the supply chain

The MTO/MTS supply chain coupling model of the above two-stage supply system is based on one premise. That is, before all kinds of materials arrive, part of the material has the technological basis of advanced production. Similarly, the quantity and type of each order are not consumed in the first place, which provides a management basis for the manufacturer's batch delivery

3.5.3. MTO/MTS Supply Chain Coupling Model

(1) *Supply Chain Coupling Model Based on Customer Order Separation Point.* Time-dimension Customer Order Decoupling Point (TCODP) is used to describe the Time course from the Customer's Order to the customized product delivered to the end customer. It divides the whole supply chain into two parts: the order pull chain contributes (S2 chain), and the inventory push chain contributes (S1). This section takes the MTO/MTS supply chain with a short life cycle order as the research object, and adopts the link analysis method to decompose the supply chain into a process consisting of a series of five interactive links, including order processing, product design, procurement/supply, production and processing, and product distribution. See Table 2.

(2) *MTO/MTS Supply Chain Coupling Model Considering Economic Value.* Decoupling Point (DP) refers to the demarcation Point in the MTO/MTS supply chain. The choice of Decoupling Point plays a key role in the rapid response of consumer demand and the reduction of supply chain costs. According to the analysis in the MTO/MTS supply chain
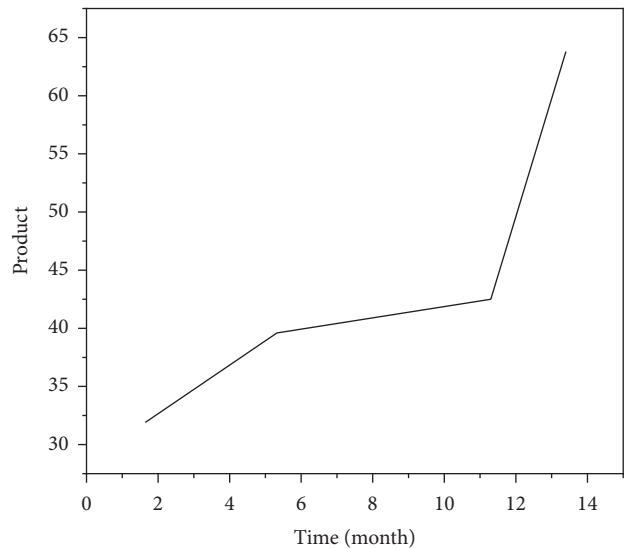
FIGURE 6: Classification of products (broken lines).
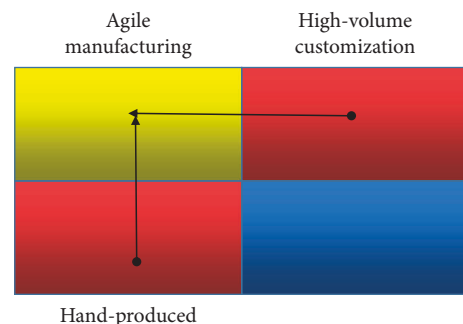
FIGURE 7: MTO/MTS supply chain coupling for product matching.

coupling model from the perspective of flexible and low-cost response, the left side of the decoupling point is the push production mode based on prediction, and the right side is the pull production mode based on order. This point is likely to become

Table 4: Retail channel sales indicators.

Month	Price	Seasonal factors	Promotional efforts	Industry form	The number of into the store	Sales volume
1	1299	0.562	1	0.98	4806	38
2	1499	0.835	0.8	0.94	4345	54
3	1499	1.123	0.8	0.95	4021	73
4	1499	1.062	0.8	0.9	3967	71
5	1499	1.259	0.8	0.95	3452	82
6	999	1.456	0.8	0.93	4089	95
7	999	1.366	0.8	0.88	4496	89
8	1499	1.442	0.3	0.84	4372	95
9	1499	1.275	0.3	0.74	3751	85
10	1499	1.318	0.3	0.78	4154	88
11	1299	1.382	0.2	0.76	4248	91
12	1299	0.941	0.1	0.74	3908	61
13	1299	0.742	0.1	0.74	3628	48
14	1299	0.623	0.1	0.67	3288	42

Table 5: Various indicators of online channel sales.

Month	Price	Seasonal factors	Promotional efforts	Industry form	The number of into the store	Sales volume
1	1299	0.389	1	0.99	0.98	577
2	1499	0.695	0.9	0.95	0.97	1023
3	1499	1.225	0.9	0.97	0.94	1815
4	1499	1.198	0.9	0.91	0.89	1772
5	1499	1.412	0.9	0.95	0.89	2088
6	999	1.938	1	0.94	0.84	2865
7	999	1.885	1	0.88	0.88	2783
8	1499	1.956	0.5	0.86	0.91	2897
9	1499	1.339	0.5	0.76	0.82	1978
10	1499	1.075	0.5	0.78	0.78	1588
11	1299	1.208	0.3	0.78	0.81	1759
12	1299	1.108	0.3	0.76	0.75	1639
13	1299	0.677	0.3	0.76	0.69	999
14	1299	0.242	0.3	0.69	0.72	575

Table 6: Parameter values of retail channel fitting sales curve equation.

	m	p	q	β	γ
Retail channel foundation Bass model	1148.88	0.05	0.24	—	—
Retail channel improvement in the Bass model	798.52	0.06	0.28	0.03	1.18

Table 7: Comparison of data forecasting effect indicators between retail channels and online channels.

Evaluating indicator	Improved the bass retail channels	The SVM retail channel	Evaluating indicator	Improved the bass network channels	The SVM network channel
RMSE	4.7442	1.5316	RMSE	92.2839	32.0485

the point of delayed manufacturing and order separation, and also the supply chain inventory control point. The economic comparison between forecast-driven and order-driven production is shown in Table 3:

(3) Economic Value Analysis Dimension Analysis. This section will analyze the influencing factors of Decoupling Point from three dimensions: material versatility, customer demand diversity, and delivery time. Material versatility: refers to the extent to which materials can be used for different purposes and by different users, depending on the practicability of the product and the position of the Decoupling Point. When the material can meet the needs of multiple types of customers or when there is no market segmentation, its versatility is high; with the further processing of materials from raw materials to components (semifinished products) and then, to finished products, the universality of materials is gradually reduced.
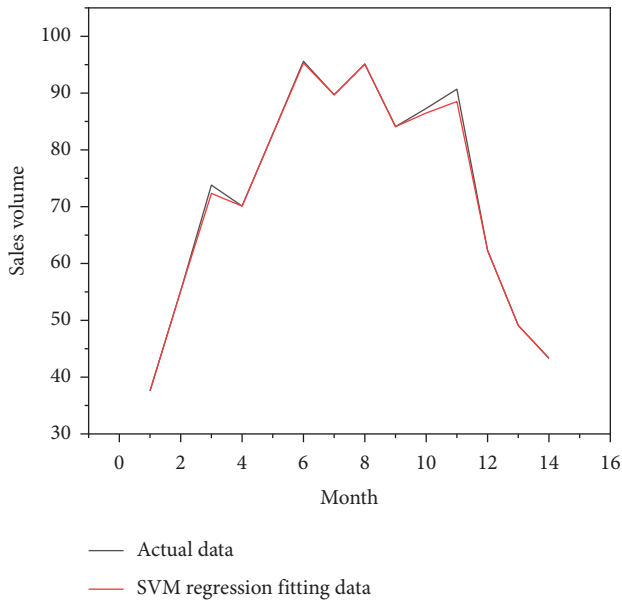
FIGURE 8: Data of SVM model retail channel sales.

FIGURE 9: Data diagram of SVM model network channel sales fitting.

(4) *MTO/MTS Supply Chain Coupling Model Based on Short Life Cycle Products*. By jumping out of the limitations of traditional functional products and innovative products and just comparing different products, the classification of the same product under different demand-deterministic environments is expanded and expanded; the nature of the product is similar to functional products when its demand is determined and innovative products when its demand is uncertain. See Figure 6.

Mass procurement production and agile response are both designed to meet the different needs of consumers.

FIGURE 10: Sales forecast data diagram of SVM model retail channel.

They both produce customized products in the mode of flexible and low-cost response but adopt different technical routes. The maximum profit model of the supply chain is used to explain the response. They are the best choice to meet consumer demand under certain constraints. See Figure 7.

4. Results Analysis

Suppose a manufacturer sells a class of fashionable electronic products with a high replacement rate through dual channels. Among them, one channel is the traditional retail channel, and the other channel is to directly conduct transactions with end customers through its own e-commerce outlets. The sales volume and influencing factor indicators of retail channels and online channels are shown in Tables 4 and 5.

According to the basic Bass model constructed in Section 2 and the programming of the improved Bass model, combined with the optimization method, Matlab2010 is used to perform nonlinear fitting to improve the goodness of fit. First, the 14-month sales data of the retail channel are fitted, and the parameter values of the fitting curve equation can be obtained, as shown in Table 6.

As can be seen from Table 6, compared with the basic model of retail channels, the potential purchase volume of retail channels under the improved Bass model decreases slightly, and the coefficient of innovation group and imitation group increase slightly, which more accurately reflects the market demand of retail channels.

Put the data of the training group of retail channels and network channels into the SVM model for calculation, and put the data of the test group into the SVM model for prediction, and you get the comparison chart of the simulated sales volume and the actual sales volume of the SVM model for the first 14-months of retail channels and network
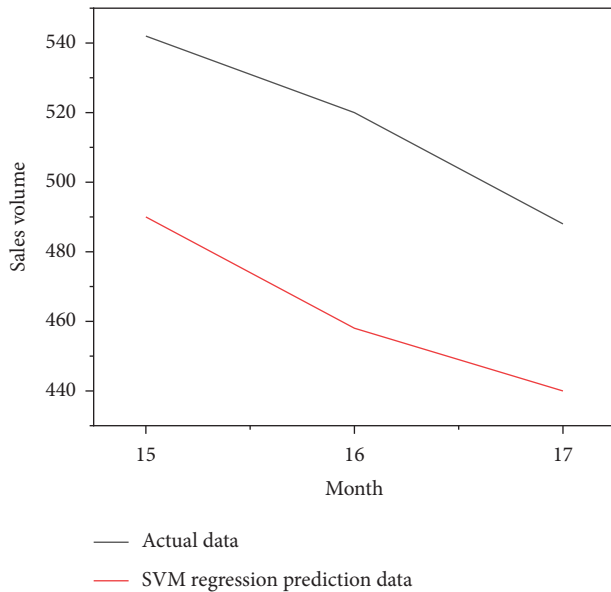
FIGURE 11: Sales forecast data of SVM model.

channels, as shown in Figures 8 and 9, as well as the forecast sales and actual sales in the next three months, as shown in Figures 10 and 11.

Table 7 further compress the prediction effect index RMSE, that is, the regression root mean square, or the fitting standard deviation of the regression system. It can be seen from the table that, whether it is a retail channel or a network channel, the RMSE value of the effect index predicted by SVM is smaller than the RMSE value of the improved Bass. It can be seen that the SVM demand forecasting model constructed by comprehensively considering multiple input factors can obtain a more accurate forecasting effect.

According to the definition of dual channels mentioned above, whether it is traditional retail channels or online channels, the short life cycle products sold originate from manufacturers and ultimately flow to end customers. Obviously, the total market demand forecast is composed of the superposition of demand forecasts of two channels, that is, the forecasted sales volume from the dual-channel system.

5. Conclusions

Based on the work practice and literature research and according to the double channel supply chain ordering mode and the characteristics of short life cycle products, a comprehensive ordering optimization method and system are proposed. This method system can not only effectively solve the existing problems of short life cycle products in dual-channel supply chain ordering, but also enrich the content of dual-channel supply chain theory. The demand forecasting model is optimized and the demand forecasting model under the dual-channel supply chain is proposed. Traditional demand forecasting methods, such as quantitative forecasting method, the time series method, and causal analysis, assume that the law of demand in the past will remain unchanged in the future. With the rapid iteration of products with a short life cycle and the rapid change of

consumer demand in the e-commerce environment, the accuracy of demand prediction is difficult to guarantee by traditional forecasting methods. In the existing literature, some scholars reduce the impact of demand uncertainty and improve the accuracy of demand prediction from the perspectives of repeated purchase, demand diffusion, increased order times, and demand renewal, respectively. However, the above studies did not consider the demand impact factors of products with short life cycle and the demand overlap and conflict brought by e-commerce dual channels. To this end, the article puts forward the characteristics of considering the influencing factors of product demand, the increase of demand information channels of dual-channel supply chain, the superposition effect of O_2O demand interference and conflict, the change of O2O demand alone or coupled change, and deduces the demand prediction model of optimization of product demand factors and the O2O demand prediction model of dual-channel supply chain. At the same time, products are classified from the perspective of demand certainty, which provides a theoretical basis for the study of the MTO/MTS coupling orders.

Data Availability

The data are available upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Dobroszek, M. Biernacki, and M. Macuda, "Management accounting in logistics and supply chain management: evidence from Poland," *Zeszyty Teoretyczne Rachunkowości*, vol. 162, no. 106, pp. 153–175, 2020.

[2] J. H. Yang, "A study on the reasonable choice and utilization of incoterms 2020 rules from the perspective of logistics and supply chain management," *Journal of Korea Trade*, vol. 25, no. 1, pp. 152–168, 2021.

[3] N. Nautiyal, S. Bisht, B. Joshi, N. Bisht, and H. Pandey, "Blockchain design for logistics and supply chain management in developing regions," *International Journal of Recent Technology and Engineering*, vol. 8, no. 6, pp. 1711–1716, 2020.

[4] I. M. Yakymchuk, "Organizational and managerial innovations in increasing efficiency of logistics using the example of a supply chain," *The Problems of Economy*, vol. 2, no. 44, pp. 325–330, 2020.

[5] Z. Zhang, X. Yang, Y. Xue, L. Gong, and J. Li, "Supply Chain Logistics Information Quality Evaluation from Just-In-Time Perspective," *IEEE Access*, vol. 8, p. 1, 2020.

[6] Q. Yaping and M. Bossman, "Logistics and supply chain management efficiency strategy for Ghana's mining industry," *European Journal of Business and Management Research*, vol. 6, no. 2, pp. 193–202, 2021.

[7] C. Di, "Research on the product logistics cost control strategy based on the multi-source supply chain theory," *Intelligent Automation and Soft Computing*, vol. 26, no. 3, pp. 557–567, 2020.

[8] B. Musigmann, H. Gracht, and E. Hartmann, "Blockchain Technology in Logistics and Supply Chain Management—A

Bibliometric Literature Review from 2016 to January 2020," *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1–20, 2020.

[9] J. Bruzda, "Multistep quantile forecasts for supply chain and logistics operations: bootstrapping, the garch model and quantile regression based approaches," *Central European Journal of Operations Research*, vol. 28, no. 1, pp. 309–336, 2020.

[10] C. Paciarotti and F. Torregiani, "The logistics of the short food supply chain: a literature review," *Sustainable Production and Consumption*, vol. 26, no. 2, pp. 428–442, 2021.

[11] S. Shivankar and R. Deivanathan, "Product design change propagation in automotive supply chain considering product life cycle," *CIRP Journal of Manufacturing Science and Technology*, vol. 35, no. 1, pp. 390–399, 2021.

[12] N. Hackius and M. Petersen, "Translating high hopes into tangible benefits: how incumbents in supply chain and logistics approach blockchain," *IEEE Access*, vol. 8, no. 1, pp. 34993–35003, 2020.

[13] H. Zarei and M. Rasti-Barzoki, "Mathematical programming and three metaheuristic algorithms for a bi-objective supply chain scheduling problem," *Neural Computing & Applications*, vol. 31, no. 12, pp. 9073–9093, 2019.

[14] J. Leukel and V. Sugumaran, "How novice analysts understand supply chain process models: an experimental study of using diagrams and texts," *Journal of Enterprise Information Management*, vol. 35, no. 3, pp. 757–773, 2022.

[15] Z. Wang, X. He, H. Shen, S. Fan, and Y. Zeng, "Multi-source information fusion to identify water supply pipe leakage based on SVM and VMD," *Information Processing & Management*, vol. 59, no. 2, p. 102819, 2022.

[16] X. Shu, J. Shen, G. Li, Y. Zhang, Z. Chen, and Y. Liu, "A flexible state-of-health prediction scheme for lithium-ion battery packs with long short-term memory network and transfer learning," *IEEE Transactions on Transportation Electrification*, vol. 7, pp. 2238–2248, 2021.

[17] J. Zhang, Z. Li, R. Tan, and C. Liu, "Design and application of electronic rehabilitation medical record (ermr) sharing scheme based on blockchain technology," *BioMed Research International*, vol. 2021, no. 2, pp. 1–12, 2021.

[18] A. A. Taleizadeh and M. Noori-Daryan, "Pricing, manufacturing and inventory policies for raw material in a three-level supply chain," *International Journal of Systems Science*, vol. 47, no. 4, pp. 919–931, 2016.

[19] A. Barman, R. Das, and P. K. De, "An analysis of optimal pricing strategy and inventory scheduling policy for a non-instantaneous deteriorating item in a two-layer supply chain," *Applied Intelligence*, vol. 52, no. 4, pp. 4626–4650, 2022.

[20] M. Aboytes-Ojeda, K. K. Castillo-Villar, and S. D. Eksioglu, "Modeling and optimization of biomass quality variability for decision support systems in biomass supply chains," *Annals of Operations Research*, vol. 314, no. 2, pp. 319–346, 2022.

Research Article

Forest Environmental Carrying Capacity Based on Deep Learning

Song Linshu ⓘ,[1] Wang Hao ⓘ,[1] Yang Chao ⓘ,[1] Song Weiming ⓘ,[1] Wang Siyi ⓘ,[2] and Wang Shen[2]

[1]College of Economics and Management, Beijing Forestry University, Beijing 100083, China
[2]Chinese Research Academy of Environmental Sciences, Beijing 100012, China

Correspondence should be addressed to Yang Chao; wanghao2022@bjfu.edu.cn and Wang Siyi; wang_si_yi@126.com

In this paper, we proposed an assessment system of forest environmental carrying capacity from many aspects and comprehensively evaluated and predicted the forest environmental carrying capacity of 40 cities in the Yangtze River Delta of China by using the proposed deep learning-based model. In addition, the proposed model is used to dynamically evaluate the comprehensive scores of forest environmental carrying capacity of 34 provinces and cities in China between 2015 and 2020. This paper also has the following contributions: (1) a deeply integrated unidirectional bidirectional LSTM considering the correlation of time series is proposed. The bidirectional LSTM network is used to deal with the backward dependence of input data to prevent the omission of some useful information, which is beneficial to the prediction results. (2) In terms of missing data processing, the Mask layer is added to subtly skip the processing of missing data, which will help to enhance the evaluation results.

1. Background

The decline of ecological environment quality directly threatens the quality of life and health of human beings [1, 2]. The term "health," which was first used in medicine to describe the life state and behavior of organic individuals, was adopted and developed by ecologists and environmentalists in their respective fields. For example, several concepts have been proposed by some scholars: environmental health, environmental medicine, forest ecosystem health, wetland ecosystem health, watershed health, landscape health, etc. [3]. Forest ecosystem health has been accepted and used more and more by ecologists, forestry, and natural resource management scholars, and "forest health" has been regarded as the standard and goal of forest status evaluation and forest resource management, and the assessment of forest ecosystem health has been put forward. After the 1990s, a series of monitoring programs have been initiated around the world [4] such as EMAP (environmental monitoring and assessment program) and FHM (forest health monitoring). It is used to understand forest state and assess forest ecosystem health, opening up a

research approach to explore sustainable research, development, and utilization of forest resources.

Ecological carrying capacity is an important basis for evaluating regional sustainable development, as well as an important indicator to measure the degree of human social and economic activities on the utilization and interference of natural resources [5]. As a component of the terrestrial ecosystem with the largest area and highest biomass, forest carrying capacity will directly affect global ecosystem security and human survival and development [6]. Therefore, the study of forest ecological carrying capacity is of great significance for coordinating the relationship between regional forest resources and sustainable economic and social development.

Since American scholars, Parker and Burgess, first defined carrying capacity as the maximum number of an individual living creature in a specific environment in 1921 [7], many foreign scholars have discussed the connotation and research methods of carrying capacity. Meadows from MIT constructed the famous "World Model" [8] by using the system dynamics model in The Limits to Growth, making carrying capacity research the core hot issue in this field for the first time. Then, Slesser began to apply the ECCO model

to the system dynamics model [9]. Daily and Ehlrich [10] focused on the relationship between population and Earth carrying capacity. Saveriades [11] proposed the concept of tourism carrying capacity for the first time in their study of ecological carrying capacity on the east coast of Cyprus.

Zhang et al. [12] believed that carrying capacity theory is a philosophical issue related to the future fate of mankind and will always be a hot and difficult issue in ecological research. Liu and Zhu [13] discussed the comprehensive, regional, and human-land relationship of geography in the application of resource and environmental carrying capacity. By definition, Gu [14] thinks that ecological carrying capacity is to point to in a certain time, within the scope of a certain space, the ecological system under the action of self-adjusting and positive human healthy and orderly development, and the ecological system can support the resource consumption and environmental pollution degree, the intensity of social and economic development, and the population of a certain level of consumption. Huang and Ren [15] studied the relationship between ecological carrying capacity and ecological security. In terms of research content, Yang and Yang [16] and Fu et al. [17] have evaluated the natural resources and environmental carrying capacity of Gansu province and Hainan Province from the provincial perspective, respectively. Jiao et al. [18], and Yu and Sun [19] studied water ecological carrying capacity and land ecological carrying capacity, respectively, from the perspective of resource types. In terms of research methods, Li et al. [20] conducted a comprehensive evaluation of Marine ecological carrying capacity in Shandong Province by using the artificial neural network method. Ma and Wang [21] used system dynamics to simulate and predict the ecological carrying capacity of Dalian city. Tian et al. [22] studied the ecological carrying capacity of Hubei Province by using the ecological footprint model.

2. Index System of Forest Ecosystem Health Assessment

Forest ecosystem health evaluation is mainly based on the stability and sustainability of forest ecosystem and the completeness (integration) of ecosystem structure and function. The stability, sustainability, and integration of ecosystem are the foundation of ecosystem health and the standard of forest ecosystem health evaluation. Only when the structure of an ecosystem is complete and the system is relatively stable, can it fully realize its ecological process and ecological function, and maintain the sustainability of the system, such an ecosystem is a healthy ecosystem. Evaluation factors include ecosystem vitality, organizational structure, resistance, and resilience.

Vigor refers to the energy input and nutrient cycling capacity of ecosystem, and its specific indicators are primary productivity and material cycling of ecosystem. Resilience refers to a system's ability to overcome stress and bounce back when the stress is gone. Organization refers to the complexity of the system, which will change and function with the secondary succession of the ecosystem.

The forest vigor can be measured by net primary productivity (NPP) of forests, biomass, and metabolism. NPP is

mainly measured by experiment and investigation, biomass is mainly calculated by investigation and model, and metabolism can be measured by biological methods.

Organizational structure refers to the species composition structure of the system and the relationship between species, reflecting the complexity of ecosystem structure. The organizational structure of ecosystem includes two meanings: one is the species diversity of ecosystem, and the other is the complexity of ecosystem. The meaning of species diversity includes both the number of existing species and the relative abundance of species.

Forest adaptation includes resilience, which refers to the system's ability to gradually recover in the absence of external stress, and resistance, which refers to the system's ability to resist external disturbance. It is difficult to measure resilience and resistance directly.

The resistance and resilience of forests are generally measured by indirect means. In forest health evaluation, the degree of forest pests and diseases in the study area or forest fire risk level can be selected to measure.

3. Determination of Biomass

Table 1 shows the evaluation index system in this article, and the following section in section 3 will introduce those indicators for details.

3.1. Sample Wood Standard. The standard tree in the ground for each wood inspection chooses the most close to the average diameter of the DBH 1 tree as a sample wood, and the selection of sample wood should avoid the following conditions: trunk serious bending, trunk section serious deformation, double branches, no roof, disease rot, the main shoot, or large branch dead.

3.2. Determination of Tree Layer Biomass. On the basis of standard field investigation, the relative growth method is adopted to calculate the biomass of tree layer. The relative growth method is the use of DBH square by tree height (D^2H) as the independent variable of fitting model, it can more truly reflect the biomass with diameter at breast height (D) and the change trend of height (H), and the calculating formula is $W = a(D^2H)^b$ or $LnW = Lna + bLn(D^2H)$, so through the survey of standards of D and H, we can work out standard ground biomass, and then calculating the biomass per hectare.

In standard plots, DBH (D) and height (H) of tree species were measured and recorded. The selected sample wood should be measured tree height, DBH, and crown; marked at 1.3 m on the ground; and marked on the trunk of the north and south orientation. After finishing the retest, we cut down the sample wood and measure the height of the live branch and the diameter of the trunk at that height immediately. We measure the height and azimuth of branches one by one from bottom to top and number them one by one. All branches were sawed together to separate branches and trunks, and then, the biomass of trunks, branches, leaves, and roots was measured, respectively.

TABLE 1: The evaluation index system.

		Sample wood standard
		Biomass of trunk
		Biomass of branch
	Determination of tree layer biomass	Determination of leaf biomass
Determination of biomass		Biomass of root
	Determination of biomass of shrub layer	
	Determination of biomass of herbaceous layer	
	Determination of plant diversity index	
	Classification of forest fire risk levels	

(1) Biomass of trunk: after removing the trunk of the branch, the disc was cut in sections from the trunk base to the top tip, with specific heights of 0 m, 1 m, 1.3 m, 3 m, 5 m, and 7 m. The last tip of the old tip should be reserved and weighed. The sample tree number, the disk number, the height of the section, and the north-south orientation were marked on the nonworking surface. Then, the diameters of each ring on the disk section were measured in the north-south and east-west vertical directions according to the requirements of the tree science. The volume of wood and bark in each section was calculated according to the skin diameter, peeling diameter, and dividing section length. Then, the bark and wood were separated, and their volumes were measured, respectively. After being dried in an oven at 85°C for 48 h, the bark was taken out and weighed to calculate the volume weight of wood and bark. Then, according to the volume of wood and bark in the segment and their volume weight, the dry matter weight of wood and bark in the segment was obtained. The total dry matter weight of wood and bark per tree was obtained by accumulative values of each section of the same sample tree.

(2) Biomass of branch: base diameter (D_0) and branch length (L) of all branches of the whole tree were measured branch by branch. Ten branches with different sizes were selected as samples, and the samples were marked. The length of the segment was determined according to the diameter and length of branches. Cut branches in each segment to distinguish the segment of the disc. The dry matter weight of wood and bark of sample branches was measured and calculated by the method of trunk. Add the wood and bark together to get the total weight of dry matter (W). Taking D_0^2L of the sample branch as independent variable and dry matter weight (W) of the branch as dependent variable, the regression analysis was conducted according to $W = a(D_0^2L)b$, and parameters a and b were obtained to establish the regression equation of the branch biomass. The D_0 and L of all branches of a single sample tree were substituted into the regression equation one by one to calculate the biomass of all single branches, and the biomass of sample wood was obtained by summing them up.

(3) Determination of leaf biomass: the sampled branches were selected to pick the leaves and weigh the fresh weight, and a small number of leaves were selected as samples. The fresh weight of the extracted leaf samples should be immediately weighed, and the leaf samples should be put into a drying oven immediately after being brought back indoors, and dried at a constant temperature of 85°C for 24 h. The dry matter ratio of leaves was calculated, and the dry matter weight of leaves per branch was further calculated. Then, with the dry matter weight of leaf as the dependent variable, the regression equation between the dry matter weight of leaf and D_0^2L was established by using the same method as the biomass of branches, and the dry matter weight of the whole plant was calculated by the relative growth method.

(4) Biomass of root: dig out all the root piles and roots of the sample wood as far as possible, and classify them according to their coarseness. Weigh the fresh children of root piles and roots of each coarseness and take samples to weigh their fresh weight. After drying for 48 h, the drying weight of the sample was weighed, and the dry matter ratio of each component and the dry matter weight of each root were calculated. We total the dry matter weight of root pile and root of all levels, namely, root biomass of the sample wood.

3.2.1. Determination of Biomass of Shrub Layer. Two to five representative quadrats were selected to dig out all shrubs with roots, separate branches, and roots, and weigh the fresh weight of stems, branches, leaves, and roots by species. Meanwhile, the wet weight and drying weight of different parts of each species were sampled to calculate the dry matter rate, and the biomass of each species and the total biomass of shrub layer in the quadrat were further calculated, and the total biomass of shrub layer per hectare was converted.

3.2.2. Determination of Biomass of Herbaceous Layer. Two to five representative quadrants were selected, and all herbaceous plants were cut down to the ground and weighed for fresh weight. A 0.5 m × 0.5 m quadrant was set in one corner of the quadrant, and the roots of all herbaceous plants

in the quadrant were dug out and weighed for fresh weight after the soil was removed. At the same time of weighing fresh weight, the aboveground part and root system were, respectively, weighed. After drying, the drying weight was weighed, and dry matter rate was calculated. Finally, biomass and biomass per hectare were calculated.

3.2.3. Calculation of Total Biomass.

The biomass of the whole community can be obtained by summarizing the biomass of each layer of tree, irrigation, and grass per unit area, and then, the subfunction relationship between the total biomass and DBH (D) and height (H) can be established by using formula $W = a(D^2L)b$. In this way, the biomass of the sample plot can be calculated by investigating D and H of the sample plot, and then, the biomass of each hectare can be calculated.

3.3. Determination of Plant Diversity Index.

The following factors should be considered in the selection of forest plant diversity indicators:

(1) The selected plant diversity index is simple and easy to operate.

(2) The selected index must be a standard quantity at a certain spatial scale, which should not only quantitatively represent the relative size of plant diversity in a specific community but also be comparable in analyzing the relative size of plant diversity among forest communities.

(3) The selected plant diversity index should reflect the differentiation of plant diversity of selected community species and each structural community as much as possible, and reflect the spatial differentiation characteristics of plant diversity quantity in the community to a certain extent.

(4) The characteristics of specific community data should be considered, which should not only meet the variable requirements in the diversity index formula but also have clear ecological significance.

According to the above principles, research needs, and collected data, the Simpson diversity index was selected in this study to represent the forest plant diversity. Simpson diversity index, as an analysis index, reflects the plant diversity of the community. This diversity index can basically be employed to handle the spatial distribution and change characteristics of the plant diversity of the forest community, and its calculation formula is as follows:

$$S = \frac{N(N-1)}{\sum_{i=1}^{m} n_i (n_i - 1)}, \tag{1}$$

where S is Simpson diversity index of the overall species in the forest community, N is the total number of individuals, n_i is the number of individuals of the ith species, and m is the number of species.

3.4. Classification of Forest Fire Risk Levels.

Forest fire is the natural enemy of forest, and it is a natural disaster with wide occurrence, great harm, strong timeliness, and difficult disposal and rescue in the world. The basic principle of forest fire risk classification is based on the "National Forest Fire Classification" as the standard, but it is different from the national classification. For forest fire danger division using clustering method, the influence of many factors, forest fire selection effect on fire bigger height, crown density and management measures, the distance from the distance from the tourist spots, advantage tree species, forest-road six factors as factor analysis, clustering to area, and 40 small class as a unit, the cluster factors of each cluster unit were investigated and assigned according to the actual survey data. Then, according to the actual experience, the SPSS statistical software was used for cluster analysis, and the threshold value was 0.8. The Forest Farm was divided into three fire risk levels:

Level I Danger Zone. The distribution area is characterized by large population density, large population mobility, close distance to the forest road, convenient transportation, railway line, and road through different site conditions of forest types of combustible materials accounted for a large proportion, which are the area of high fire risk, in the forest health evaluation value of 0.4.

Level II Danger Zone. The distribution area is characterized by close proximity to tourist spots and roads, and its stand also has obvious characteristics. The stand structure is complex, with more shrubs and dead objects in the forest, and the proportion of combustible objects in different site types is large. The value of this area in forest health evaluation is 0.4.

Level III Danger Zone. Located in the southwest corner of Forest Farm, the distribution area is characterized by relatively concentrated small classes, far from tourist spots, mostly natural broadleaved mixed forest, steep slope, high altitude, and little human interference. The value of this area in forest health evaluation is 0.8.

4. The Comprehensive Evaluation Model Construction for Forest Carrying Capacity

4.1. Recurrent Neural Network.

Cyclic neural network can well process sequence information and learn effective features from it, and is applied to image [23], speech recognition, and other fields in the early stage [24]. In recent years, RNN has also achieved good performance on tasks in the field of NLP. Lin et al. [25] used hierarchical RNN to model documents and proposed a two-step training method, which achieved good results in sentence ordering task and oral translation task. Liu et al. [26] proposed a multi-time-scale LSTM model extracts sentence and document representation. Experimental results on four data sets indicate that this model is superior to other neural network models. Zhou

et al. [27] constructed a bidirectional LSTM network based on an attention mechanism to learn features from sentences for relational classification, only considering the word vector of sentences, without considering other lexical or syntactic features generated by NLP tools. In document-level emotion classification tasks, researchers generally construct hierarchical neural network models to learn document feature representation. For example, Tang et al. [28] first used the LSTM network to learn sentence vector representation from word vector and then used bidirectional threshold RNN (gated RNN) to learn document vector representation from sentence vector. Chen et al. [29] constructed a two-layer LSTM network and considered the attention mechanism of users and products.

RNN can not only learn the information in the sentence but also learn useful feature representation from the syntax tree path. Roth and Lapata [30] used the LSTM network to extract features of dependency paths for semantic role labeling. In terms of the task of relation classification, Xu et al. [31] adopted multichannel LSTM network and considered four dependency paths of words, parts of speech, dependency relation, and hypernym among entities. Cai et al. [32] proposed a RNN network, which contains bidirectional two-channel LSTM network, learning word path, and dependency path are two characteristic representations of dependency path. Similarly, the proposed model includes a bidirectional LSTM network for learning valid features from syntactic paths.

4.2. LSTM Network Structure. The LSTM network structure was first proposed by Hochreiter and Schmidhuber in 1997 [33]. Figure 1 shows the structure diagram of an LSTM basic unit. Like RNN, LSTM elements have layer input x_t and layer output h_t at each iteration. The complex unit also takes into account the unit input state $\widetilde{C}_t$, the unit output state C_t, and the previous unit output state C_{t-1} when training and updating parameters.

In LSTM network, the basic structure of the hidden layer is a storage unit, and the structure shown in Figure 2 is an LSTM cell structure unit. It contains one or more memory cells and a pair of adaptive, multiplicative connectivity cells that connect inputs and outputs to all cells in the block. At the core of each memory cell, it is a circular self-connecting linear cell, commonly known as "constant error transmission," and the activation of this cell is called "cell state," which is the key to LSTM. The constant error transmission of a unit is shown in the dotted box in Figure 2.

Constant error transmission solves the vanishing error problem: in the absence of new input or error signals, the local error return of constant error transmission remains unchanged, neither growing nor decaying. However, if only constant error conveyor belt, LSTM structure is unable to realize the addition and deletion of information flow. Therefore, a similar gating structure is proposed to realize the screening in the process of information transmission. As shown in Figure 3, the gated structure cleverly filters the information, mainly through a neural layer of sigmoid function and a point-by-point multiplication operation.
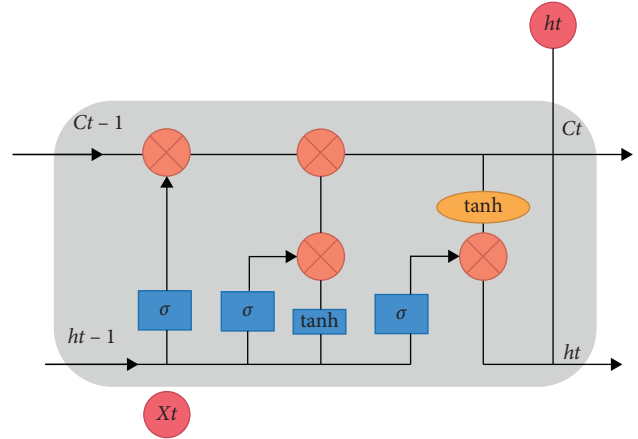
Figure 1: The architecture of LSTM.
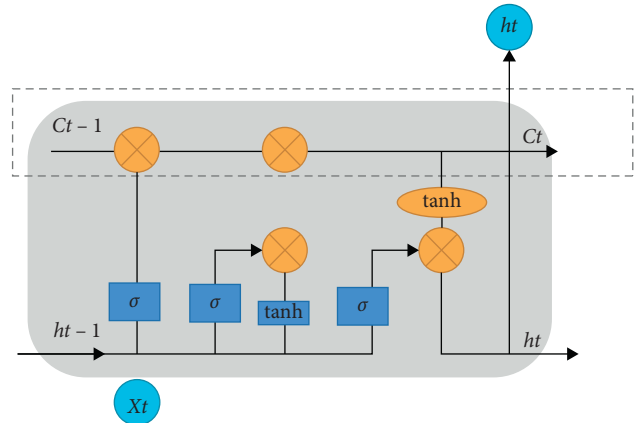
Figure 2: Constant error transmission structure.
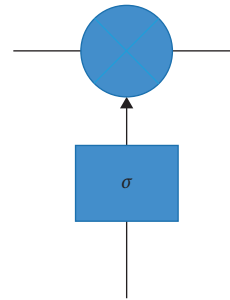
Figure 3: Gate control structure.

Through the sigmoid layer structure, an output value between 0 and 1 can be obtained. These output values imply a weight assignment of passing information (e.g., 0 means that no information is allowed to pass through the structure, and 1 means that all input information can pass through the structure).

In summary, LSTM unit includes 3-gate structures. These gated structures, especially forgetting gates, help LSTM become an efficient and extensible model for solving several learning problems related to sequential data. For the t th time step, the forgetting gate, input gate, and output gate are expressed as f_t, i_t, and o_t, respectively.

The first step we need to consider in the LSTM network unit is what redundant information the network structure wants to discard that we do not care about. This decision requires the forgetting gate structure to operate. The forgetting gate will receive the input x_t of this unit and the output h_{t-1} of the hidden layer of the previous time step as the input of the cell unit at this moment, and learn according to the following formula (2). It learns a real value between 0 and 1 for each item in C_{t-1}, so as to control the degree of forgetting of the state of the previous unit:

$$f_t = \sigma_g\big(W_f x_t + U_f h_{t-1} + b_f\big), \tag{2}$$

where W_f is the weight matrix that maps the hidden layer input to the forgetting gate, U_f is the weight matrix that connects the output state of the previous unit to the forgetting gate, and b_f is the bias vector. σ_g is the gate activation function. Usually, the hyperbolic tangent function (tanh) is used as the activation function in LSTM structures.

The next step is to decide how much new information to add to the cellular state. An input gate and a hyperbolic tangent function (tanh) control what new information is added. Then, according to the output f_t of the forgetting gate obtained previously, it is used to control the extent to which the information of the previous unit is forgotten, and the output i_t of the input gate is used to control how much new information is used. Then, we can update the unit state of the memory unit according to the following formulas:

$$i_t = \sigma_g\big(W_i x_t + U_i h_{t-1} + b_i\big), \tag{3}$$

$$\tilde{C}_t = \tanh\big(W_C x_t + U_C h_{t-1} + b_C\big). \tag{4}$$

Similarly, W_i and W_C are weight matrices that map hidden layer input to legacy input gate and input unit state, respectively, while U_i and U_C are weight matrices that connect previous unit output state to input gate and input unit state, respectively. b_i and b_C are bias vectors.

Finally, we need to make sure that our network produces the desired output, which will be based on our cell state. The output gate is employed to consider how much information the network needs for further learning. Similar to the updating of the two parts of the input gate, the final output content of the output gate is also determined by activation function, as shown in the formula (4). The tanh function is then used to process the content of cell state (formula (5)). Each computed Ct value may not be within the range of -1 to 1, so it needs to be adjusted. We multiply these two pieces and get the part that we want to output.

$$o_t = \sigma_g\big(W_o x_t + U_o h_{t-1} + b_o\big), \tag{5}$$

$$h_t = o_t * \tanh\ (C_t). \tag{6}$$

Similarly, W_o is the weight matrix that maps the hidden layer input to the output gate, while U_o is the weight matrix that connects the output state of the previous unit to the output gate. b_o is the bias vector. According to the results of the four equations above, during each iteration, unit output state C_t and layer output h_t can be calculated as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \tag{7}$$

Finally, the C_t obtained is the cell state of the hidden layer unit.

The final output of the LSTM layer should be a vector of all outputs, represented by $Y_T = [h_{T-n}, \cdots, h_{T-1}]$.

4.3. Two-Way LSTM Network Prediction Model.
Bidirectional propagation LSTM network (Bi-LSTM) is also a deformation of bidirectional cyclic network structure. So, let us first introduce bidirectional cyclic networks (BiRNN).

4.3.1. BiRNN Network. Since the classical RNN usually include the time step lag among the input and the target when handling the sequence problem, some context information can be given to the network; that is, the future information of T step time step can be added to obtain the output together. Theoretically, T can be very large to capture all future available information, but in fact, it is found that if T is too large, the accuracy of prediction results may decline. This is because the network concentrates on learning and memorizing a large amount of input information and does not reasonably learn the input information, which leads to the decline of the joint modeling ability of the prediction model.

BiRNN can be modeled by processing forward and backward data propagation. Such bidirectional transmission not only provides historical knowledge for each neuron node in the network but also imperceptibly learns the trend of subsequent changes.

In the forward propagation layer, a forward hidden layer output h_t^f is obtained through a nonlinear transformation of input data according to the formula (8). In the backward propagation layer, a nonlinear transformation is performed on the input data according to the formula (8) to obtain a backward hidden layer output h_t^b.

$$h_t^f = \sigma\big(W_{hh}^f h_{t-1}^f + W_{xh}^f x_t + b_f\big), \tag{8}$$

$$h_t^b = \sigma\big(W_{hh}^b h_{t+1}^b + W_{xh}^b x_t + b_b\big). \tag{9}$$

Two recurrent neural network (RNN) hidden layers are trained, and the two independent hidden layers pass through a connection layer. According to the formula (10), the output of forward transfer layer and the output of reverse transfer layer are combined into the same output layer.

$$h_t = W^f h_t^f + W^b h_t^b + b_h. \tag{10}$$

4.4. Structural Optimization of the Prediction Model.
For the network framework we have built, we need to refine our predictions. Predecessors have tried a variety of methods to optimize the model as a whole. In our article, the dropout and early-stop algorithms are employed to optimize the proposed framework to promote the Acc of the proposed method.

TABLE 2: Forest health grade division criterion.

Health grade	I	II	III	IV	V
Network output value	0	1	2	3	4
Health condition	Healthy	General healthy	Sub-healthy	Unhealthy	Other lands

TABLE 3: Forest health condition of the cities.

Health condition	Small class number
Healthy	7, 37
General healthy	1, 3, 9, 12, 28, 33, 40, 41, 4, 5, 8, 10, 13, 14, 16, 22, 23, 25,
Subhealthy	26, 27, 30, 32, 35, 38, 39
Unhealthy	11, 18, 19, 20, 36
Other lands	21, 29

4.5. Dropout Principle. Hinton proposed a method called dropout, which can effectively reduce the occurrence of overfitting, which is equivalent to adding a regularization constraint to the model. This method will be employed in this paper.

5. Experiment and Result Obtained

In order to unify standards and ensure the reliability of experimental results, it is generally necessary to preprocess the original data before modeling, which is commonly referred to as standardized data processing. We need to transform the original data into dimensionless standardized values to eliminate the influence caused by different attributes among different indicators, so as to make the results more comparable. Otherwise, the accuracy of experimental results may be affected to some extent.

$$X' = \frac{X - X_{\max}}{X_{\max} - X_{\min}}, \tag{11}$$

$$X = 2 * X' - 1. \tag{12}$$

$X_{\max}$ is the maximum value in the sample instance, and $X_{\min}$ is the minimum value in the sample instance. The normalized sample data X' are scaled to a value between -1 and 1 through the formula (12).

5.1. Standardization of Factors. As the evaluation factors of the evaluation index system come from different aspects, it is difficult to directly evaluate them, because the dimensionality between the coefficients is not unified, so there is no comparison. Even for the same parameters, although their impact on health can be judged according to the size of their measured values, there is still a lack of a comparable environmental standard to accurately reflect their contribution to health. Therefore, the evaluation factors must be quantified and the incomparability of parameters can be solved by standardized methods. The standardization method is as follows:

$$X_j = \frac{X_i}{Y_{\max}}, \tag{13}$$

where X_j is standardized index value; X_i is the value of the selected index; and $Y_{\max}$ the highest value in a class of indexes.

5.2. Forest Ecosystem Health Result

5.2.1. Classification of Forest Health Status. In the process of forest health evaluation, this paper chooses BP neural network model to evaluate the forest farm. According to its evaluation method and selected index system, on the basis of systematic analysis and integration of existing research results at home and abroad, the forest ecosystem health evaluation method based on Bi-LSTM is constructed. The grading standards of health are indicated in Table 2.

5.2.2. Prediction of Forest Carrying Capacity. This paper takes 40 urban forests in the Yangtze River Delta of China as research samples. Table 3 shows the forest health condition of the cities.

As can be seen from Table 3, there are 7 subhealth areas in carrying capacity, accounting for 17% of the total number of major areas. There are 18 districts, or 45 per cent of the main districts, with average health. There were 5 undesirable areas, accounting for 12.5% of the total number of major areas. Only two districts, or 5 percent of the total, are healthy.

5.2.3. Forest Carrying Capacity in Yangtze River Delta from 2015 to 2020. Figures 4 and 5 show the Local Moran scatter diagram of high-quality forest development of urban agglomerations in the Yangtze River Delta in 2015 and 2020 under the adjacency matrix. It could be found from the figure that Moran's I index is positive and most cities are concentrated in the first and third quadrant, namely, high and low agglomeration areas, indicating that the forest development level of cities in the Yangtze River Delta has a positive spatial correlation.

5.2.4. Forest Carrying Capacity in China (Part of the Area) from 2015 to 2020. According to Figures 6 and 7, the comprehensive evaluation scores of forest environmental carrying capacity of 34 provinces and cities in China
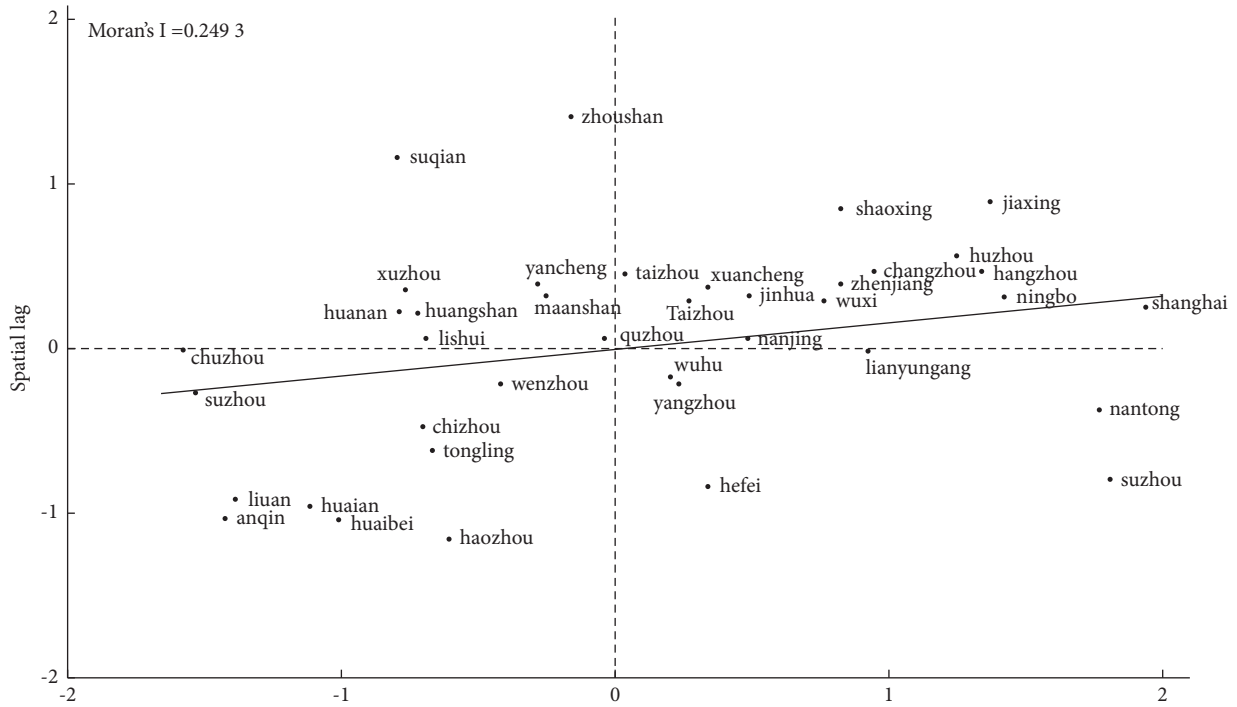
Figure 4: Standardized value of forest development level in 2015.
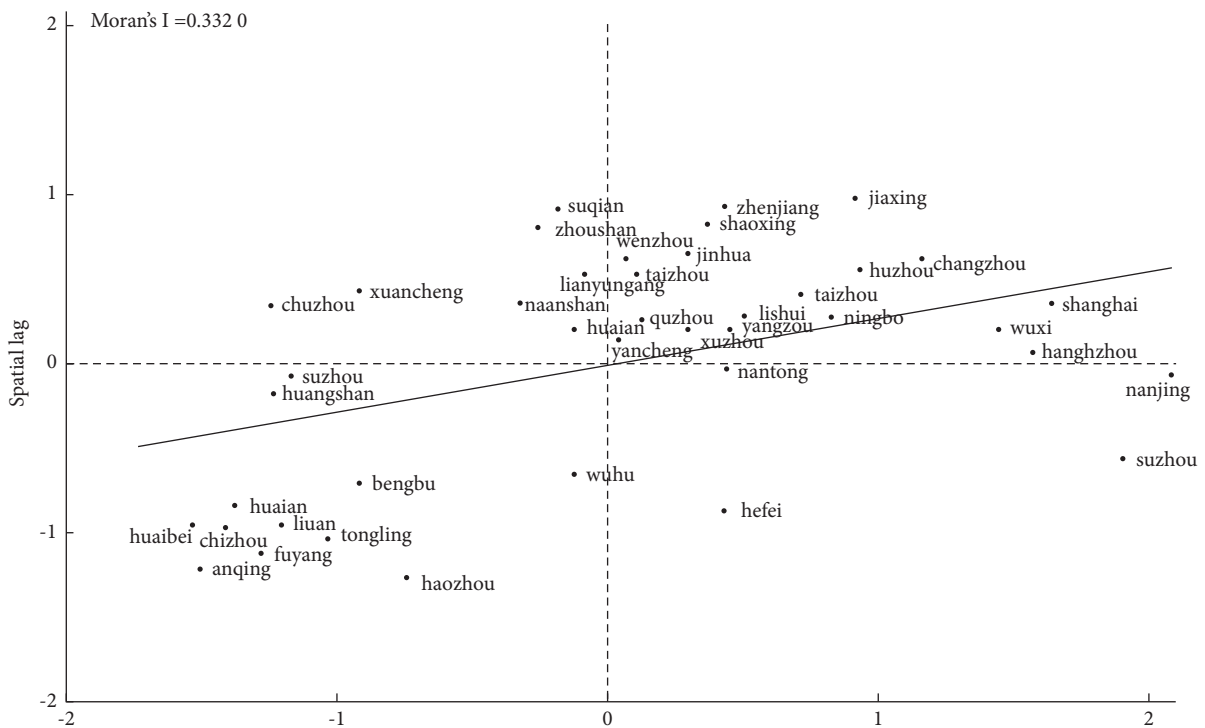
Figure 5: Standardized value of forest development level in 2020.

changed from 2015 to 2020. It can be found that: (1) the forest environmental carrying capacity of Tibet, Xinjiang, and Inner Mongolia has always maintained a high level; (2) compared with 2015, the forest environmental carrying capacity of almost all cities in 2020 has been improved, which indicates that the intensity of environmental protection and forest protection in China has been increasing.

Year 2020

☐	≤ 0.24	☐	0.58~0.75
☐	0.24~0.41	☐	> 0.75
☐	0.41~0.58		
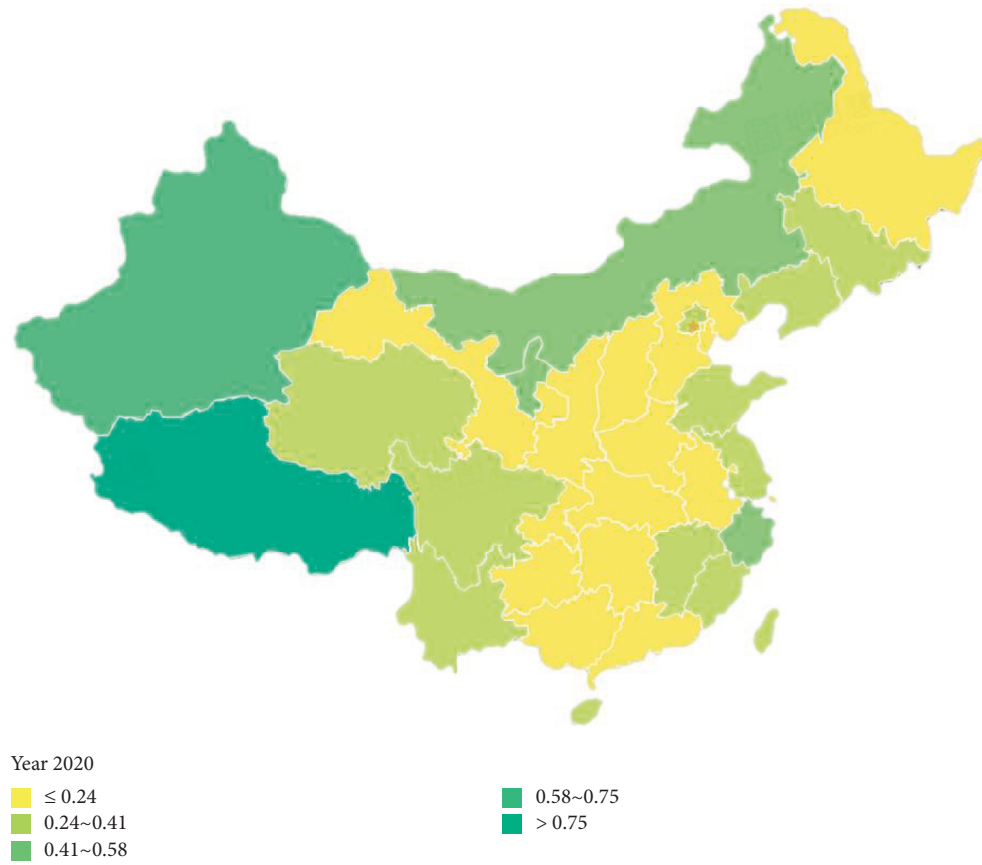
FIGURE 6: Comprehensive evaluation value of forest environmental carrying capacity (2020 year).

Year 2015

☐	≤ 0.22	☐	0.52~0.67
☐	0.22~0.37	☐	> 0.67
☐	0.37~0.52		
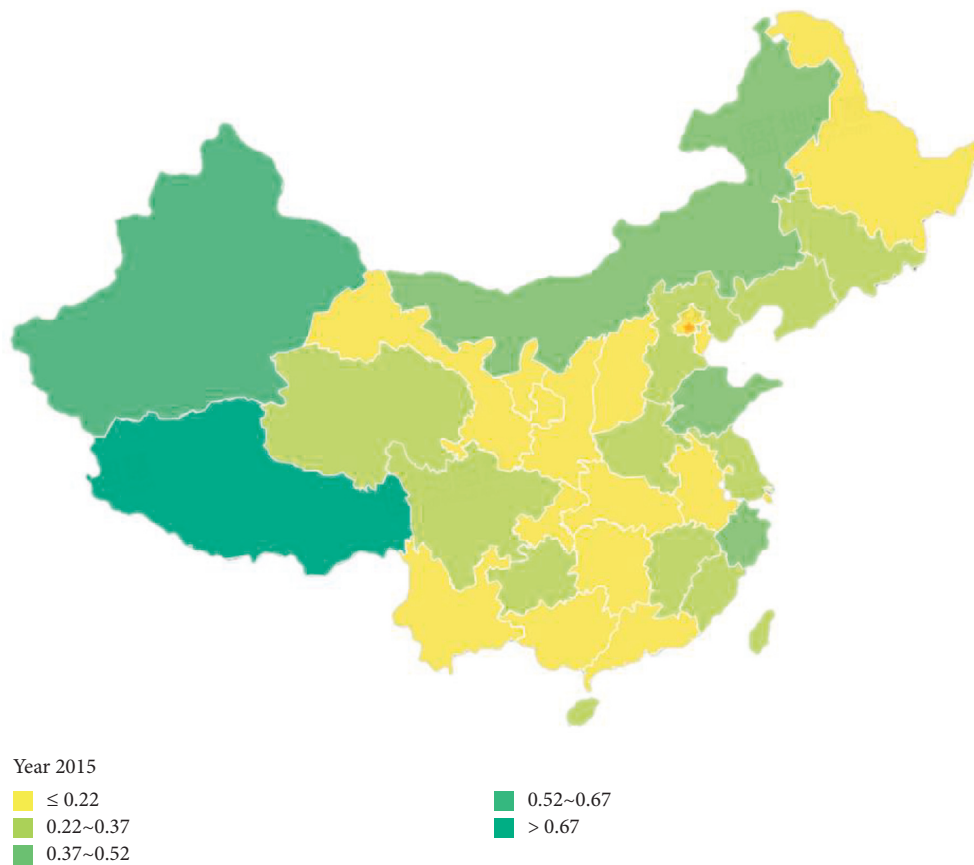
FIGURE 7: Comprehensive evaluation value of forest environmental carrying capacity (2015 year).

6. Conclusion

The improvements and contributions of this study mainly focus on the following aspects: (1) a deeply integrated unidirectional bidirectional LSTM considering the correlation of time series is proposed. The bidirectional LSTM network is used to deal with the backward dependence of input data to prevent the omission of some useful information, which is beneficial to the prediction results. (2) In terms of missing data processing, the Mask layer is added to subtly skip the processing of missing data. The results indicate that the proposed method in this paper could predict and evaluate the comprehensive index of forest environmental carrying capacity well. On the basis of this study, further improvement and expansion can be made. A neural network based on graph structure can be added to the model to learn and interpret spatial features.

Data Availability

Data for this article can be obtained by contacting the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Linshu Song and Hao Wang contributed equally to this work.

Acknowledgments

References

[1] J. Oszlanyi, "Forest health and environmental pollution in Slovakia," *Environmental Pollution*, vol. 98, no. 3, pp. 389–392, 1997.

[2] A. Szepesi, "Forest health status in Hungary," *Environmental Pollution*, vol. 98, no. 3, pp. 393–398, 1997.

[3] K. A. Blatner, C. E. Keegan, J. OLaughlin, and D. L. Adams, *Forest health management policy:A case study in southwestern Idaho Assessing Forest Ecosystem Health in the Inland West*, R. N. Sampson, Ed., pp. 317–338, Food Products Press, New York, NY, U.S.A, 1993.

[4] S. A. Alexander and C. J. Palmer, "Forest health monitoring in the United Stales:First four years," *Environmental Monitoring and Assessment*, vol. 55, no. 2, pp. 267–277, 1999.

[5] J. X. Gao, *Theory Explore of Sustainable Development: Ecological Carrying Capacity Theory,method and Application*, China Environment Science Press, Beijing, China, 2001.

[6] E. Ritter and D. Dauksta, "Human–forest relationships: ancient values in modern perspectives," *Environment, Development and Sustainability*, vol. 15, no. 3, pp. 645–662, 2013.

[7] X. R. Guo, X. Q. Mao, and S. H. Ran, "Research progress in environmental carrying capacity in China," *China Population, Resources and Environment*, vol. 10, no. S1, pp. 29–31, 2000.

[8] D. Meadows, *The Limits of Growth*, pp. 15–268, The Commercial Press, Beijing, China, 1984.

[9] M. Slesser, *Enhancement of carrying capacity options-ECCO: simulation software for assessing national sustainable development The Management of Greed*, Resource Use Institute, China, 1990.

[10] G. C. Daily and P. R. Ehrlich, *Population, Sustainability, and Earth's Carrying Capacity Ecosystem Management*, pp. 435–450, Springer, New York, NY, U.S.A, 1994.

[11] A. Saveriades, "Establishing the social tourism carrying capacity for the tourist resorts of the east coast of the Republic of Cyprus," *Tourism Management*, vol. 21, no. 2, pp. 147–156, 2000.

[12] L. B. Zhang, W. H. Li, X. F. Liu, and W. Wang, "Carrying capacity: origin, development and prospective," *Acta Ecologica Sinica*, vol. 29, no. 2, pp. 878–888, 2009.

[13] W. Z. Liu and J. Zhu, "Research progress of resources and environmental carrying capacity: from the perspective of the comprehensive study of geography," *China Population, Resources and Environment*, vol. 27, no. 6, pp. 75–86, 2017.

[14] K. K. Gu, "The concept of ecological carrying capacity and its research methods," *Ecology and Environmental Sciences*, vol. 21, no. 2, pp. 389–396, 2012.

[15] Q. Huang and Z. Y. Ren, "Ecological carrying capacity and ecological security," *Journal of Arid Land Resources & Environment*, vol. 18, no. 2, pp. 11–17, 2004.

[16] L. J. Yang and Y. C. Yang, "Study on the spatiotemporal variation in resource environmental carrying capacity in the Gansu province of China," *Acta Ecologica Sinica*, vol. 37, no. 20, pp. 7000–7017, 2017.

[17] G. J. Fu, H. L. Xu, and W. T. Chen, "Natural ecological carrying capacity in Hainan Province," *Journal of Natural Resources*, vol. 23, no. 3, pp. 412–421, 2008.

[18] W. J. Jiao, Q. W. Min, and W. H. Li, "Measuring water ecological carrying capacity with the ecosystem-service-based ecological footprint (ESEF) method: an application in Huzhou City in the Tai Lake basin," *Resources and Environment in the Yangtze Basin*, vol. 25, no. 1, pp. 147–155, 2016.

[19] G. H. Yu and C. Z. Sun, "Land carrying capacity spatio-temporal differentiation in the bohai sea coastal areas," *Acta Ecologica Sinica*, vol. 35, no. 14, pp. 4860–4870, 2015.

[20] M. Li, S. Y. Dong, and H. H. Zhang, "Studies on the early warning and evaluation of the carrying capacity of sea area in Shandong Province based on the multi-dimensional state space and neural network model," *Marine Science Bulletin*, vol. 34, no. 6, pp. 608–615, 2015.

[21] Z. Q. Ma and L. Wang, "Simulation and prediction of ecological bearing capacity based on system dynamics in urbanization process of Dalian City," *Journal of Central South University of Forestry & Technology*, vol. 33, no. 4, pp. 70–75, 2013.

[22] L. L. Tian, J. Luo, Y. Dong, H. T. Liu, and J. X. Zeng, "Temporal and spatial dynamic research of the ecological footprint and ecological carrying capacity of Hubei Province," *Resources and Environment in the Yangtze Basin*, vol. 25, no. 2, pp. 316–325, 2016.

[23] R. Cierniak, "A new approach to image reconstruction from projections using a recurrent neural network," *International Journal of Applied Mathematics and Computer Science*, vol. 18, no. 2, pp. 147–157, 2008.

[24] C. F. Juang, C. T. Chiou, and H. J. Huang, "Noisy speech recognition by hierarchical recurrent neural fuzzy networks," in *Proceedings of the 2005 IEEE International Symposium on Circuits and Systems*, pp. 5122–5125, IEEE, Japan, May 2005.

[25] R. Lin, S. Liu, M. Yang, M. Li, M. Zhou, and S. Li, "Hierarchical recurrent neural network for document modeling," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 899–907, Portugal, September 2015.

[26] P. Liu, X. Qiu, X. Chen, W. Shiyu, and H. Xuanjing, "Multi-timescale long short-term memory neural network for modelling sentences and documents," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, pp. 2326–2335, Portugal, September 2015.

[27] P. Zhou, W. Shi, J. Tian et al., "Attention-based bidirectional long short-term memory networks for relation classification," in *Proceedings of the 54th annual meeting of the association for computational linguistics*, vol. 2, pp. 207–212, Germany, August 2016.

[28] D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, pp. 1422–1432, Portugal, September 2015.

[29] H. Chen, M. Sun, C. Tu, Y. Lin, and Z. Liu, "Neural sentiment classification with user and product attention," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, pp. 1650–1659, Austin, TX, U.S.A, November 2016.

[30] M. Roth and M. Lapata, "Neural semantic role labeling with dependency path embeddings," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Germany, August 2016.

[31] Y. Xu, L. Mou, G. Li, C. Yunchuan, P. Hao, and J. Zhie, "Classifying relations via long short term memory networks along shortest dependency paths," in *Proceedings of the 2015 conference on empirical methods in natural language processing*, pp. 1785–1794, Portugal, September 2015.

[32] R. Cai, X. Zhang, and H. Wang, "Bidirectional recurrent convolutional neural network for relation classification," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 756–765, Germany, August 2016.

[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

Retraction

Retracted: Research and Development of User Clustering-Based Content Similarity Algorithms in Dance-Assisted Choreography Techniques

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Wu and M. Liu, "Research and Development of User Clustering-Based Content Similarity Algorithms in Dance-Assisted Choreography Techniques," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1364835, 10 pages, 2022.

Research Article

Research and Development of User Clustering-Based Content Similarity Algorithms in Dance-Assisted Choreography Techniques

Yanyan Wu [1] and Min Liu[2]

[1]*School of Architecture and Art, Central South University, Changsha 410000, Hunan, China*
[2]*College of Music and Dance, Huaihua University, Huaihua 418000, Hunan, China*

Correspondence should be addressed to Yanyan Wu; 206134@csu.edu.cn

With the gradual development of digital information and software computing capabilities, the use of computers in dance-assisted choreography is becoming more and more widespread. But although the level of computers is now in rapid development, the technical level of using computers in dance choreography is not yet very mature, technical support is not in place, dance-assisted choreography is not effective, and the existing technical level is not yet able to meet the new needs of dance choreography. In order to improve the dance-assisted choreography technology and provide a more complete educational user interface for dance-assisted choreography, the content similarity algorithm of user clustering has a wide range of operations and a strong ability to calculate the amount of data, combined with the computer to apply the content similarity algorithm of user clustering in dance-assisted choreography technology to build a dance-assisted choreography system based on user clustering. The article proposes three major methods based on collaborative filtering algorithm of user clustering, collaborative filtering algorithm based on similarity class and user preference, and fuzzy cluster analysis of users and analyses their principles. In the experimental part, the performance of IBCF algorithm and collaborative filtering algorithm in dance-assisted choreography system is compared and analysed to observe the change of MAE value under the change of user similarity with number under different k values of cluster classes. The experimental results found that the MAE values of the IBCF algorithm and the collaborative filtering algorithm in the system were at 0.84 and 0.76, respectively, with a difference of about 8% between the two MAE values. The smaller the MAE value, the higher the effectiveness in the dance-assisted choreography technique. Applying the clustering algorithm to the system to make local adjustments and analysis of dance movement paths, it can grasp the choreography rules more precisely and innovate the choreography techniques.

1. Introduction

With the development of science and technology, the use of computers has become more widespread and the development trend of combining computing technology for dance choreography is getting better and better. In the process of dance-assisted choreography, how to meet the different needs of different people for dance choreography under the development of the times, improve the dance-assisted choreography system, analyse the movements for the given movement fragments, and apply the analysis in the new

dance-assisted choreography technology has become a problem that more and more people need to pay attention to. This paper investigates the effect of the content similarity algorithm in the development of a dance-assisted choreography system by combining it with user clustering. Experimental analysis is conducted for the proposed algorithm, and the experimental results find that the collaborative filtering algorithm has obvious advantages in the dance-assisted choreography system, its overall MAE value is smaller than that of the IBCF algorithm, and the MAE of the IBCF algorithm and the collaborative filtering algorithm is

0.84 and 0.76, respectively, when the analysis data set is selected at 40, with a difference of about 8% between the two MAE values; as the number of k values increases, the values on the recommendation results as well as the accuracy and recall of the collaborative filtering algorithm increased, confirming the high effectiveness of the collaborative filtering algorithm in the dance-assisted choreography system; compared to the traditional dance-assisted choreography system, the results of the analysis of the index of each factor of the dance-assisted choreography system under the user clustering algorithm found that the effectiveness of the dance choreography rose by 15%. The dance-assisted choreography system with various user clustering algorithms was found to be more effective in capturing human movement to some extent.

The analysis of user clustering-based content similarity algorithms in dance-assisted choreography technology can be developed to meet the different needs of dance-assisted choreography, and with the support of the algorithms, the dance-assisted choreography system is made better and more adaptable to the current development of the web. With algorithms such as content similarity for user clustering, dance movements can be analysed in depth to improve the movement path of movement data, master the essence of dance-assisted choreography, and facilitate the improvement of dance technology.

Driven by social developments, the use of cluster analysis is becoming more and more widespread. Virmani et al. [1] aimed to analyze the role of relational clustering behind group interactions in social networks. Clustering enhances the predictability between users and the ability to discover like-mindedness. They use integrated K-means clustering to extract entities and their corresponding interests based on skills and location by aggregating user profiles across multiple online social networks. The proposed integrated clustering utilises known K-means algorithms to improve the results of aggregating user profiles across multiple social networks. It turns out that good integrated clusters can be generated to envision the discoverability of users for specific interests [1]. Bouras and Tsogkas [2] propose a new approach to personalised recommendations that combines user and text clustering with other information retrieval (IR) techniques based on the algorithm they developed, W-k-means, in order to provide users with articles that match their profiles. Experimental results show that by aggregating items and user clustering using multiple IR techniques (e.g., classification and summarisation), the recommender generates better results than when using each or both of these techniques but without applying clustering [2]. Chandler et al. [3] used statistical cluster analysis algorithms to identify natural groupings based on AE profiles in a data-driven exploratory analysis. Clusters were clinically assessed to identify clusters associated with current safety issues. Using proportional reporting rates, headache and dizziness combined with fatigue or syncope were more common in HPV vaccine reports compared to non-HPV vaccine reports in women aged 9–25 years. Cluster analysis revealed additional reports of AE following HPV vaccination that were severe in nature, describing symptoms that overlapped with those reported in recent safety signals (POTS, CRP, and CFS) but without a clear diagnosis. Although the causal relationship between HPV vaccination and these adverse events remains uncertain, a more extensive analysis of spontaneous reports could better identify relevant case series for a thorough signal assessment [3]. Ilmarinen et al. [4] used data from baseline (diagnosis) and 12-year follow-up to determine the phenotype of asthma exacerbations in adults. K-mean cluster analysis was performed to identify phenotypes using variables from 171 patients at baseline examination and follow-up. Five clusters were identified. At follow-up, these patients were on the lowest dose of ICS, but 56% were well controlled. The results of the study can be used to predict the prognosis of adult asthma patients and contribute to the development of personalised treatment [4].

Enache [5] argues that the essence of spiritual and cultural life is the artistic integration of the performer-audience division, which goes hand in hand with the subdivision of "dance" into work and play. Language is expressed in formulas, and gestures and dance are rhythmic. It is worth noting that the extent to which dance as an organic necessity can be further developed into a closed structure still needs to be analysed and studied by researchers in the field [5]. These frames offer the experience of viewing the solo from multiple perspectives. It interweaves the individual with the voices of other artists, scholars, and students to discover that they all have the ability to articulate some of the most subtle, practical, and logical aspects of their experiences in creating new works [6]. Poutanen et al. [7] report on a study that used two visions created by Microsoft as a starting point for describing a user-centred design approach that extends the ideas of the dance method to interaction design and demonstrates how micro-motion analysis can be carried out in practice. They use the structural reorganisation of the movement continuum initially presented in the video to personalise the description of the dance for the first time as a way of understanding the kinesthetic quality and potential of the implied dance. The approach emphasises the impact of interaction design on the moving and experiencing body, and the potential of the moving and experiencing body for interaction design [7]. The above analysis of the clustering algorithm and the development of the choreography is relatively clear and provides informativeness for the following depiction.

2. User Clustering-Based Content Similarity Algorithm in Dance-Assisted Choreography Research and Development Method

2.1. Collaborative Filtering Algorithm Based on User Clustering

2.1.1. User Similarity Measure. Collaborative filtering algorithms based on user clustering first use statistical search methods to find similar users in a user group, and then evaluate specific users based on the results of similar users' evaluations of specific entries. If some users rate some items similarly, they rate other items similarly, a condition that must be met for the user-based collaborative filtering

algorithm to hold. The method based on collaborative user filtering is widely used in existing recommendation systems because of its low computational effort and high accuracy. The key to this approach is to identify the neighbourhood users of the target user based on the similarity of the item scores and to feed their predicted score results back to the user. The principle of the collaborative filtering algorithm based on user clustering is shown in Figure 1.

User-based collaborative filtering algorithms have the advantage of being simple to compute and highly accurate. Collaborative filtering algorithms based on users generally include several major steps such as user similarity metrics, finding nearest neighbour users and predicting the ratings of target users. Currently, three methods, namely, cosine similarity measure, modified cosine similarity measure method, and association similarity measure method, are commonly used as user similarity measures [8].

Cosine similarity: the users' ratings of items in the user-item matrix are regarded as vectors on a multidimensional space, with the rating value set to zero at the position of the default rating, and the similarity between users can be obtained from the cosine angle of their rating vectors of items [9]. Assuming that there are two different users, a and $\underline{b}$, whose corresponding vectors in the space vector are $\vec{a}$ and $\vec{b}$, then the similarity between users a and b can be calculated as follows:

$$sim_{ab} = \cos{(\vec{a}, \vec{b})} = \frac{\vec{a}, \vec{b}}{\|\vec{a}\|\|\vec{b}\|}. \tag{1}$$

Modified cosine similarity: this is ignored in the cosine similarity measure because different users have different evaluation criteria. The average ratings $\overline{K}_a$ and $\overline{K}_b$ for users a and b can be known from the ratings, and finally, the similarity between user a and user b is calculated, giving

$$sim_{ab} = \frac{\sum_{g \in h_{ab}} (K_{a\varphi} - \overline{K}_a)(K_{b\varphi} - \overline{K}_b)}{\sqrt{\sum_{g \in h_a} (K_{a\varphi} - \overline{K}_a)^2} \sqrt{\sum_{g \in h_b} (K_{b\varphi} - \overline{K}_b)^2}}. \tag{2}$$

The correlation similarity is based on an improvement of the idea of the Pearson correlation coefficient (PCC), which, like the modified cosine similarity measure, corrects for the problem of the scale on which items are rated by users. The formula for calculating the similarity of users using correlated similarity is

$$sim_{ab} = \frac{\sum_{g \in h_{ab}} (K_{a,g} - \overline{K}_a)(K_{b\varphi} - \overline{K}_b)}{\sqrt{\sum_{g \in h_{ab}} (K_{a,g} - \overline{K}_a)^2} \sqrt{\sum_{g \in h_{ab}} (K_{b,g} - \overline{K}_b)^2}}. \tag{3}$$

Once the similarity between users is obtained through the similarity metric, the top k nearest neighbours of the current user can be obtained and the item recommendation can be made [10]. However, in the similarity calculation, it can not only help users improve their personal experience to improve the user proximity of the website, but also shorten the user's product selection time to improve the user's purchasing efficiency, but the following shortcomings are common in the existing collaborative filtering algorithms.
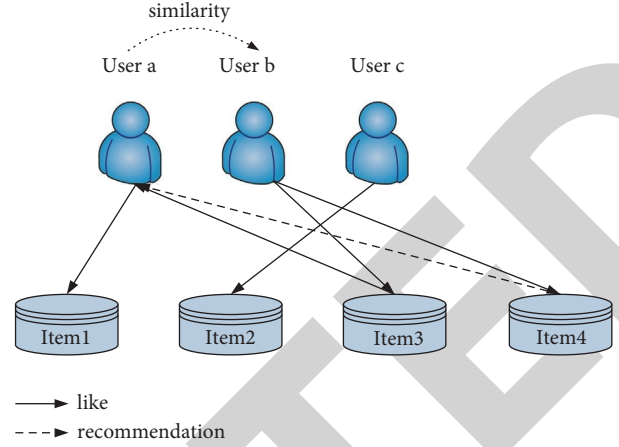
Figure 1: Recommendation principle of collaborative filtering algorithm based on user clustering.

Cold start problem: The cold start problem consists of two situations: firstly, for a new user, the system does not have any information about the user and cannot recommend any items for him/her, and secondly, for a new item, there is no information about the user's rating of the item and the system cannot recommend it to the user.

Data sparsity problem: The collaborative filtering algorithm relies on user ratings of items to make recommendations, but in most cases the user ratings are small compared to the total number of users and items. It tends to result in sparse user and item scoring matrix data with many unscored null values. In this case, it is difficult to obtain a similar set of users to the target user and to generate suitable recommendations for the target user.

Scalability issues: The collaborative filtering algorithm has to perform the work of finding nearest neighbour users after performing the similarity measure; however, when the number of users in the system is particularly large in order of magnitude, the efficiency of finding neighbour users decreases and the scalability of the algorithm is not high [11].

2.1.2. K-Means Algorithm. *K*-means is the process of dividing a collection of *m* data into specified *k* classes, where the similarity between objects is calculated using the Euclidean distance as the criterion, and the clustering is based on the principle of least squared error [12]. The operation is repeated in the clustering process to derive the optimal clusters corresponding to the initial cluster centre distribution. The final number of clusters, *k*, is specified from the given *m* data samples, and the corresponding number of cluster centroids is selected. The distance from each data object in the data set to these *k* centres is calculated, and the data are classified into clusters accordingly. After the classification is complete, the centroids of the clusters are redetermined once again, and this completes the calculation in one iteration. The better centroid is selected again in the

first classification result, the division of the data is calculated from this centroid, and the clustering of the data is divided through several iterations until the criterion function converges. Figure 2 shows the data partitioning diagram after the national porcelain iteration.

The formula for the criterion function using *K*-means is

$$R = \sum_{i=1}^{h} \sum_{j \in g_h} \left(k_j - \beta_i\right)^2. \tag{4}$$

The contour coefficients are generally used to analyse the classification effect of the *K*-means algorithm, expressed by the formula as follows:

$$M(i) = \frac{n_i - m_i}{\max(m_i, n_i)}. \tag{5}$$

When the value of the contour coefficient is closer to 1, it means that this clustering is good and the clustering is reasonable; when the contour coefficient is close to −1, it means that similar samples do not belong to this class of clusters. When the contour coefficient is close to 0, it means that the similar samples are on the boundary of the two cluster classes.

It is indicated that the contour coefficient of a class is the mean of the corresponding values through all the points in it, expressed by the formula as follows:

$$MN = \frac{\sum_{i=1}^{k} \left(k_j - \beta_i\right)^2}{k}, \tag{6}$$

where SC is the contour coefficient of the whole data set; the larger the SC value, the more appropriate the number of clusters *k* value; *m* indicates the number of objects in the data set.

2.1.3. Recommendation Results Generated by User Clustering.
The improved query method for the target user's neighbourhood set points out that the correct neighbourhood set has a significant impact on the accuracy of the recommendation results [13]. The interests and preference tendencies of the target user can be more accurately predicted by finding the neighbourhoods closest to the target user, which are closer to the target user's interests and preferences. The improved method finds the set of neighbours most similar to itself in a smaller user class cluster space when the system searches for neighbouring users. Preprocessing of users is done by clustering using *K*-means to generate *k* clusters of user classes. The user rating information is processed into a user rating matrix, from which *k* initial cluster centroids are selected, and the distance from the user to the cluster class centre is recorded according to the different ratings of each user. The method for dividing the centroids of the cluster classes is

$$H = \sum_{i=1}^{G} \sqrt{\left(k_{1i} - k_{2i}\right)^2}. \tag{7}$$

Next, the set of neighbours is filtered in the space of class clusters to which the target user belongs, using the previous
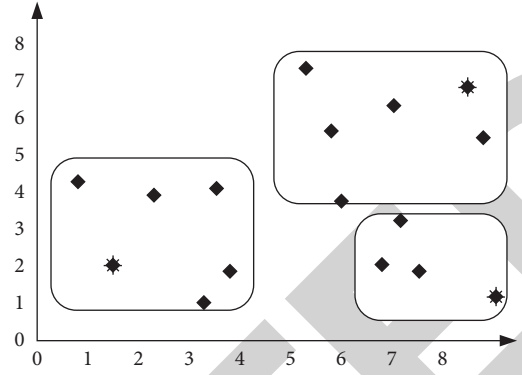
Figure 2: *K*-means algorithm after several iterations of data division.

improved research method, and the possible ratings given by the user for content not visited in that system are derived [14]. From these predicted ratings, the top rated items are selected to be recommended for the target user.

2.2. Collaborative Filtering Algorithm Based on Similar Classes and User Preferences.
In order to facilitate the subsequent selection of services based on preferences, users need to have an accurate description of their preferences in order to select the service they are most satisfied with. Due to the problem of data sparsity, it is difficult to obtain similar user sets of the target user to generate suitable item recommendations for the target user. How to overcome the sparsity problem of data in the rating matrix and accurately grasp the user's interest preferences in this limited information becomes the main concern of the recommendation system [15]. By exploring user authenticity, user preferences, and user access frequency, it recommends products that are similar to user preferences, thereby improving user ratings. The collaborative filtering algorithm based on similarity classes and user preferences focuses more on the high ratings given by users when mining their real interests, not only considering the number of visits by the user but also corresponding the user's rating information to the feature matrix of the item class.

For user preference dance moves selection, the set of similar classes and the user's preference weights are first input, and the candidate dance moves in each similar class are normalised and calculated to produce a dance move for each similar class. Firstly, the user's interest preference is initially determined from the frequency of the user's access to the feature. The feature classes belonging to the items that have been rated by the user's visits are collated and extracted, and then the user's ratings of the items are mapped to the user's ratings of the corresponding features. The number of visits to a feature category is used to determine the level of interest in that category; i.e., the ratio of the number of visits to a category to the total number of items visited by the user is the user's preference for that feature category. The initial interest of the user in the feature class in terms of the frequency of access is calculated using the formula as follows:

$$W_{a,b} = \frac{c_{a,b}}{|H_a|}. \tag{8}$$

However, the number of visits to a feature category is not sufficient to determine a user's preference for that feature category; the true preference is determined by the user's subjective rating of the feature category. It is therefore important to analyse users based on their access characteristics, so that their interest preferences are more realistic, and to calculate their true interest preferences in terms of subjective ratings.

$$Q_{a,b} = \frac{m_{a,b}}{c_{a,b}}. \tag{9}$$

Finally, the similarity of users' true preferences for feature categories is calculated based on the frequency with which they visit the feature category and the number of high ratings the user has in the feature category.

$$F_{a,b} = W_{a,b} \times Q_{a,b}. \tag{10}$$

2.3. User Fuzzy Clustering.
The user fuzzy clustering process first uses the data recorded by the computer to construct a fuzzy matrix of human motion, which is then normalised and then calibrated to produce a fuzzy similarity matrix of human motion, which is clustered and output [16]. Fuzzy clustering algorithms provide more reasonable clustering results compared to hard clustering algorithms, but increase in time complexity. In general, the methods for performing fuzzy clustering analysis are broadly divided into normalisation of sample data, calibration of the fuzzy similarity matrix, and clustering. In practice, a suitable classification should satisfy three conditions: self-reflexivity, symmetry, and transferability, and satisfying the above conditions is an equivalence relation, so fuzzy clustering analysis is based on equivalence relations. A diagram of the process of fuzzy clustering is shown in Figure 3.

2.3.1. Data Standardisation.
In fuzzy clustering analysis, the sample data should first be preprocessed. Data normalisation is the unification of each indicator value within a certain range of numerical characteristics for various numbers and orders of indicators, thus eliminating the influence of differences in units of characteristic indices and different orders of magnitude [17]. Currently, there are two common methods of data normalisation: mean-variance normalisation and great-minimum normalisation.

Given a sample set, mean-variance normalisation is calculated and the formula is expressed as

$$k_{ab}' = \frac{k_{ab} - \overline{k}_a}{s_a} \ (1 \le b \le k). \tag{11}$$

Very large and very small standardisation:

$$k_{ab}' = \frac{k_{ab} - \min\{k_{ab}\}}{\max\{k_{ab}\} - \min\{k_{ab}\}}. \tag{12}$$

2.3.2. Establishing a Fuzzy Similarity Matrix.
Calibration is to measure the degree of similarity between the classified objects, which is generally expressed by the similarity coefficient. There are many methods commonly used to calibrate similarity coefficients in fuzzy similarity matrices, such as the similarity coefficient method, the angle cosine method, the maximum-minimum method, and the arithmetic mean-minimum method. The algorithms for several calibration methods are described next.

Correlation coefficient method:

$$r_{ab} = \frac{\sum_{h=1}^n (k_{bh} - \overline{k}_b)}{\sqrt{\sum_{h=1}^n (k_{ah} - k_b)^2}\sqrt{\sum_{h=1}^n (k_{ah} - k_a)^2}}. \tag{13}$$

Clipped cosine method:

$$r_{ab} = \frac{\sum_{h=1}^n k_{ah} \cdot k_{bh}}{\sqrt{\left(\sum_h^n k_{ah}^2\right)^2 \left(\sum_h^n k_{ah}^2\right)}}. \tag{14}$$

Maximum-minimum method:

$$r_{ab} = \frac{\sum_{h=1}^n \min(k_{ah}, k_{bh})}{\sum_h^n (k_{ah} + k_{bh})}. \tag{15}$$

Arithmetic mean-minimum method:

$$r_{ab} = \frac{\sum_{h=1}^n h = 1}{1/2 \sum_{h=1}^n (k_{ah}, k_{bh})}. \tag{16}$$

2.4. User Clustering-Based Dance-Assisted Choreography System.
User clustering is applied to dance choreography to build a dance-assisted choreography system based on user clustering features. This system can capture the motion information in the data file with the support of clustering algorithms, display the captured 3D animation effect, and then provide dance path editing based on the settled animation [18]. In this system, dance movement segments and local movements and paths can be edited, and physical adjustment drills are provided so that the resulting performance movements are realistic and conform to physical laws. The choreography assumes that the performer is travelling along a straight path, meaning that the performer's path of movement is abstracted into a straight line, and editing the path of movement allows for further reuse of movement data. The implementation of the system is therefore divided into six modules: a movement file manipulation module, a 3D movement display module, a movement editing module, a user clustering module, a mathematical calculation module, and a dance clip database module [19]. A user clustering-based dance-assisted
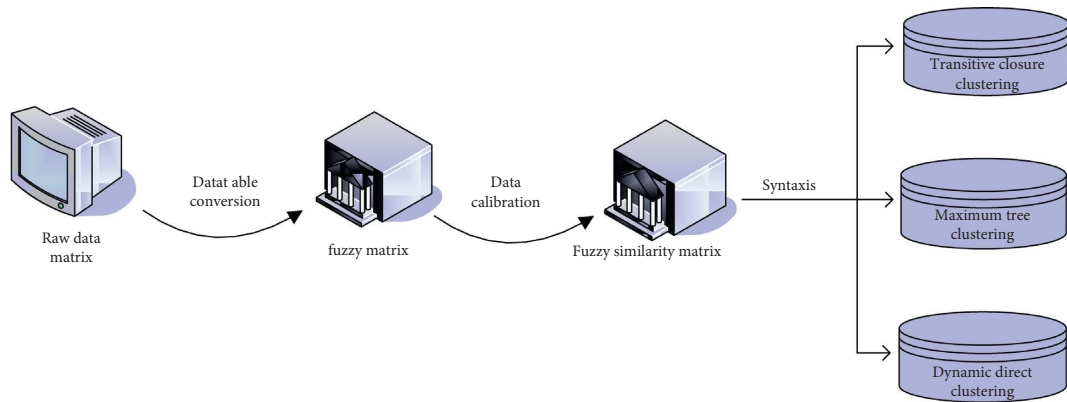
Figure 3: Fuzzy clustering composition.
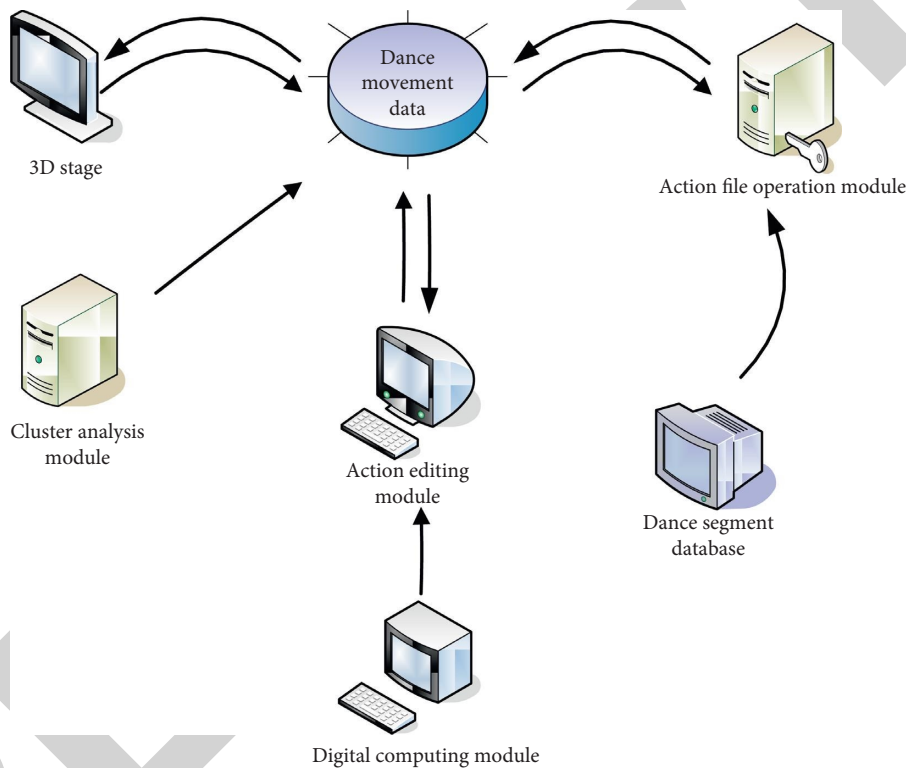
Figure 4: User clustering-based choreography system.

choreography system is shown in Figure 4. This system has an important role to play in improving and refining dance-assisted choreography techniques.

In the movement editing module, since the movement data itself only include information on the spatial position of the joints and do not contain information on the mass distribution of the body, there is also a need to focus on the impact of the physical parameters of the body on the choreography [20]. In biomechanics, the body's central nervous system automatically adjusts the angular momentum of the centre of mass to a minimum during movement, so that the centre of mass remains translational to enhance dynamic balance [21]. In the human movement system, ZMP is an important basis for determining dynamic equilibrium, when the body is in dynamic equilibrium, and

the centre of pressure of the supporting polygon is coincident. During the human balance adjustment process, the frames of the performance segments are divided into touchdown and nontouchdown categories in order to reduce the huge data calculation and to correct the human motion according to the variation of the force on the human body [22].

Touchdown frame equilibrium is when the body touches the ground and the external forces acting on it are ground support and friction in addition to gravity. After rotational filtering of the performance segment to be optimised, for each touchdown frame in that performance segment, the human pose is adjusted according to the classification of the skeletal segment set in each step, so that the adjusted ZMPs fall within the support polygon and the human body is

dynamically balanced. Nontouchdown frame equilibrium is when the body does not touch the ground and the only external force acting on it is gravity, so its centre of mass must follow a parabolic trajectory and the total angular momentum must be conserved [23]. This can be expressed in the mathematical formulation of the physical constraint as

$$\text{COM}(f) = \frac{\sum iw_f e_i(f)}{\sum iw_i},$$

$$d_{\text{COM}} = F, \tag{17}$$

where COM represents the body centre of mass position and d_{COM} represents the angular momentum of the centre of mass.

The division of the angular momentum of the centre of mass into two components can be expressed by the equation as follows:

$$d_{\text{COM}} = d_{\text{COM}}^A + d_{\text{COM}}^B. \tag{18}$$

When translations and rotations of the centre of mass of the individual limb segments of the human body take place, it is possible to obtain

$$d_{\text{COM}}^A = \sum_i g_i (e_i - COM) \times D_i,$$

$$d_{\text{COM}}^B = \sum_i K_i f_i = \sum_i K_i F_c^B \beta^i, \tag{19}$$

where F_c^B denotes the matrix in the limb section, the unit vector of the line connecting the joint points at both ends in section i.

According to the six modules in the system, the dance movements can be captured and optimised based on user clustering similarity [24]. For the captured motion, data are a collection of 3D spatial coordinates of the individual joints in the human body model over time. Think of the human body model as a tree, with the waist joint as the root node, and as the parent node moves, the child joints follow. The effect of the performance clip after optimisation is shown in Figure 5.

3. Development of a Content Similarity Algorithm Based on User Clustering in Dance-Assisted Choreography Technology

3.1. Experiments on the Application of Collaborative Filtering Algorithm in Dance-Assisted Choreography Technology. The accuracy of the algorithm for scoring predicted items is assessed by the value of the mean absolute error (MAE); the smaller the value of MAE, the better the result. As the user clustering algorithm helps to improve the quality of system recommendations, in the dance choreography system, it is assumed that there are k clusters of user dance feature classes in the system. According to the similarity calculation method, the similarity of human movements is calculated, the accuracy of the similarity algorithm is firstly experimented, and then the experimental results are used to analyse whether the class clusters can be included in the
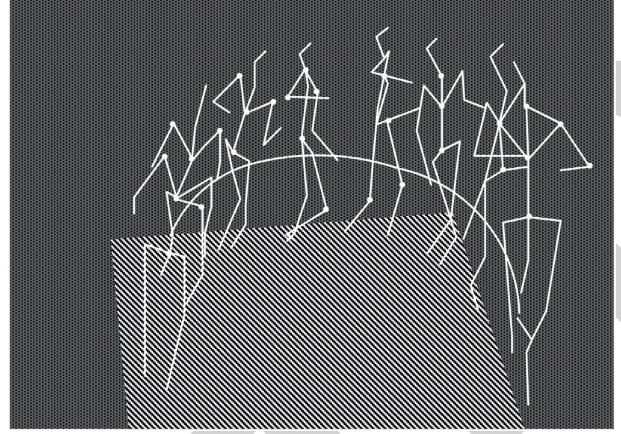
Figure 5: Analysis of dance movement segments after user clustering similarity optimisation.

dance-assisted choreography technology. To confirm the collaborative filtering algorithm, the article analyses the IBCF algorithm and the collaborative filtering algorithm under the same data situation, as shown in Figure 6 which shows the change of MAE value under the change of user similarity with number for different k values of cluster classes.

As shown in the figure, under the same conditions, the collaborative filtering algorithm proposed in the article has more obvious advantages than the IBCF algorithm, and its overall MAE value is smaller than the IBCF algorithm. When the number of actions is selected to be 40, the MAE value of the IBCF algorithm is about 0.84, and the MAE value of the collaborative filtering algorithm is about 0.76. The MAE value of the two differs by about 8%, and the collaborative filtering algorithm is used after the number of human bodies is 40. The MAE value tends to be stable, so the collaborative filtering algorithm is more effective than the IBCF algorithm.

In order to verify the effectiveness of the collaborative filtering algorithm in dance-assisted choreography and the diversity of similarity calculation, the length of the recommendation list is set to 20 in the experiment; that is, 20 items are recommended for the target user, and the collaborative filtering under different k values is analysed. The precision and recall of the recommendation list results given by the filtering algorithm are shown in Table 1.

The corresponding graph is shown in Figure 7.

The precision rate and recall rate are greater than 0.3, indicating that the algorithm recommendation quality is good. As shown in the figure, as the number of k values increases, the accuracy and recall of the recommended results of the collaborative filtering algorithm are improving. Therefore, it can be judged that the recommendation quality of the collaborative filtering algorithm is constantly improving. When the number of k values is between 10 and 50, the precision rate and recall rate increase rapidly, the two k values are above 40, and the precision rate and recall rate are both greater than 0.3. The more the number, the higher the precision rate and recall rate. The accuracy and recall rate of the recommended number after 50 tend to be stable,
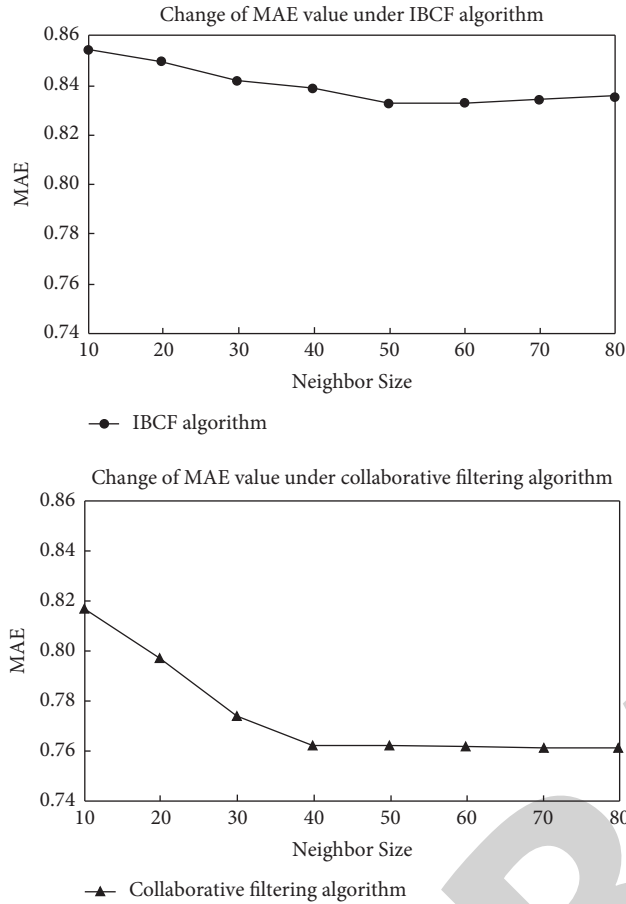
Figure 6: Comparison of MAE changes under IBCF algorithm and collaborative filtering algorithm.

Table 1: Precision and recall with different numbers of neighbours.

Characteristic class cluster k	Accuracy	Recall
10	0.1923	0.2105
20	0.2355	0.2216
30	0.3640	0.2821
40	0.4311	0.3346
50	0.4321	0.4213
60	0.4356	0.4407
70	0.4386	0.4532
80	0.4399	0.4511

indicating that the collaborative filtering algorithm is highly effective in dance-assisted choreography.

3.2. Application Experiment of User Fuzzy Clustering in Dance-Assisted Choreography Technology. Under the user fuzzy clustering analysis, for all the page sequences visited by users within a period of time, because the preprocessing part cannot get the page paths that the user is really interested in, according to the algorithm idea of this paper, the user interest matrix is intercepted with a threshold value to extract the user's sense of interest. The user fuzzy clustering analysis is applied to the dance assistant choreography system, and the speed of human movements is systematically analysed by

comparing the IBCF algorithm and the user fuzzy clustering algorithm. Judge the help of user fuzzy clustering to dance-assisted choreography technology. Figure 8 shows the time consumption of the two algorithms under the same amount of data.

Because the IBCF algorithm is more suitable for analysing objects with a small amount of data, the response to operations with a large amount of data is slow and the efficiency is low. It can be seen from the figure that as the amount of data increases, the clustering time of the two algorithms increases, the amount of data increases, and the sides become longer. In the experiment, when the maximum amount of data is 60, the time used by the IBCF algorithm is 50 s, the time used by the user fuzzy clustering algorithm is 38, and the time difference between the two algorithms is 12 s. Compared with the user fuzzy clustering algorithm, the running time of the IBCF algorithm is longer, and the user clustering effect is more obvious under the user fuzzy clustering algorithm. Under the effective clustering, over time, the dance assistant choreography system can easily adapt to user preferences, perform choreography in a targeted manner, and improve choreography techniques.

Compared with the traditional dance assistant choreography system and the index analysis of each factor of the dance assistant choreography system under the user clustering algorithm, the data information in Table 2 is obtained.

By analysing Table 2, it can be seen that in the choreography system optimised by the user clustering algorithm, the technical level of dance choreography has increased from 77% to 84%, the choreography effect has increased from 67% to 82%, the quality of choreography has increased from the original 76% to 85%, and the satisfaction with choreography has risen from 72% to 83%. The choreography effect has increased by 15%, and the room for improvement is the largest. Under the user clustering similarity algorithm, each reference factor in the dance choreography system has been improved to a certain extent.

4. Discussion

This paper researches, develops, and analyses the content similarity algorithm based on user clustering in dance-assisted choreography technology. In view of the problems and shortcomings encountered by the existing dance choreography technology, this paper uses the content similarity algorithm of user clustering. This paper proposes three methods: collaborative filtering algorithm based on user clustering, collaborative filtering algorithm based on similarity class, and user preference and user fuzzy clustering analysis. It is found that the collaborative filtering algorithm is higher than the IBCF algorithm in terms of accuracy and recall. In the application experiment of fuzzy clustering in dance assistant choreography, comparing the time consumption of IBCF algorithm and user fuzzy clustering algorithm in data analysis, it is found that user fuzzy clustering algorithm consumes less time and has better clustering effect; compared with traditional dance assistant, the index changes of each factor of the choreography system and the
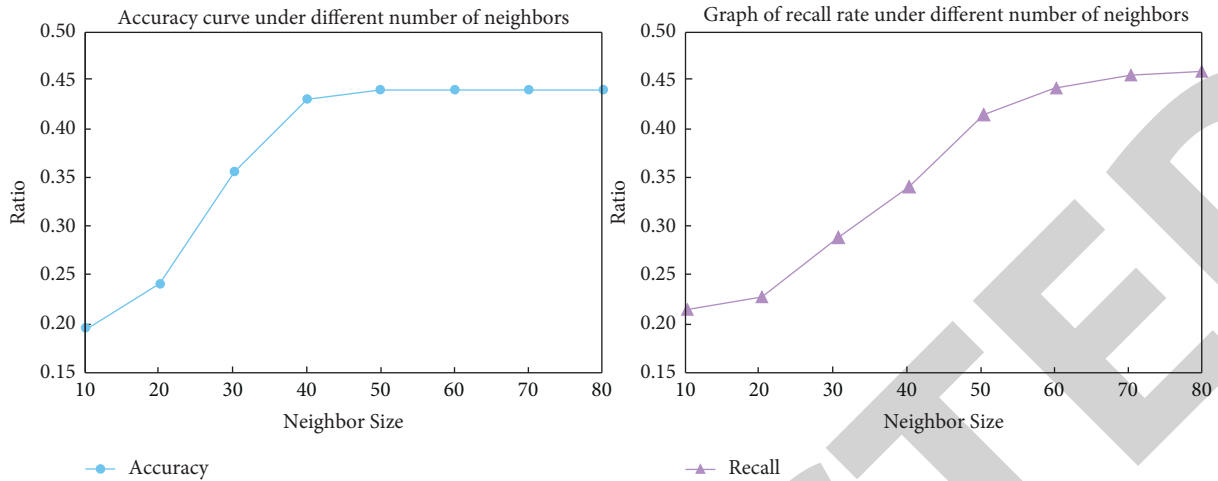
Figure 7: Plot of precision and recall for different numbers of neighbours.
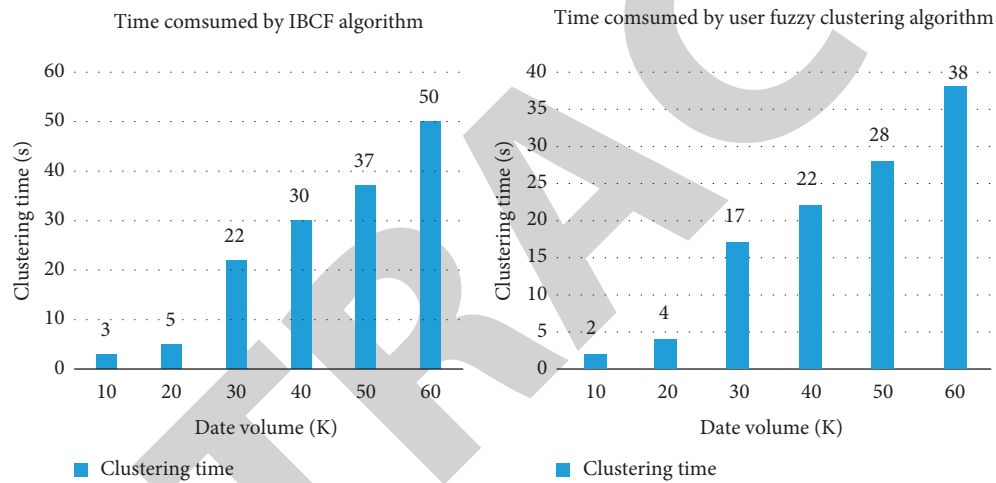
Figure 8: Time consumption comparison between IBCF algorithm and user fuzzy clustering algorithm.

Table 2: The index changes of each factor before and after algorithm optimisation.

	Traditional editing system (%)	User similarity clustering arrangement system (%)
Technical level of choreography	77	84
Choreography effect	67	82
Choreography quality	76	85
Dance choreography satisfaction	72	83

choreography system after clustering optimisation found that each reference factor was improved to a certain extent.

5. Conclusions

This paper takes user clustering as a premise and integrates several user clustering similarity algorithms to analyse dance-assisted choreography techniques. The paper proposes a collaborative filtering algorithm based on user clustering, a collaborative filtering algorithm based on similarity classes and user preferences, and a user fuzzy clustering analysis. In the collaborative filtering algorithm based on user clustering, the user similarity measure, the K-means algorithm, and the recommendation results generated by user clustering are analysed, and for the problem of data sparsity in the collaborative filtering algorithm based on user clustering, the collaborative filtering algorithm of similar classes and user preferences is proposed to mine the authenticity of household interest preferences; then, the data are standardised and a fuzzy matrix is established to the fuzzy clustering of users then generalised through data normalisation and establishment of fuzzy matrices. In the experimental part, the collaborative filtering algorithm in dance-assisted choreography technology application and user fuzzy clustering in dance-assisted choreography technology application are analysed. After the experiments, it is

Hindawi

Research Article

Research on the Performance of Network Propagation by Using the Machine Learning and Internet-of-Things Technology Integrating Model

Feng Chen [ID]

College of Artificial Intelligence, Zhejiang College of Security Technology, Wenzhou 325016, Zhejiang, China

Correspondence should be addressed to Feng Chen; chenfeng@zjcst.edu.cn

We combine machine learning with Internet of Things technology to study the performance of network propagation model. This paper first introduces the construction environment of the business push system and then realizes user clustering and active business push by using the experimental data. Experimental results show that the active service push system constructed in this paper is feasible and effective. The experiment also compares and analyzes the influence of different clustering methods on the accuracy of service push. The results show that the clustering effect of the multi-Markov chain model (m-MCM) method is superior to that of the *K*-means method, a commonly used machine learning method, and the accuracy of user-service push obtained by the m-MCM method is superior to that obtained by the *K*-means method. Finally, on the basis of the existing experimental results, the shortcomings of the service push system are summarized, the future improvement direction and specific implementation measures are proposed, and new requirements for the future update of the service push system are put forward.

1. Introduction

The industrial Internet of Things is the product of the deep integration of the Internet of Things, big data, cloud computing, and other new-generation information technologies into the manufacturing industry. It is the main starting point for the digital transformation of the manufacturing industry and has become a key support for the fourth industrial revolution. Germany, the United States, Japan, and other major developed countries have established promotion agencies for the industrial Internet of Things and formulated strategic plans for the development of advanced manufacturing industries, including the industrial Internet of Things, so as to realize the transformation and upgrading of traditional manufacturing industries and reshape their competitive advantages. China has also formulated the "Made in China 2025" policy to accelerate the in-depth integration of the new generation of information technology and manufacturing as the main line and to promote intelligent manufacturing as the main direction.

2. Related Works

The research on the industrial Internet of Things started earlier in the world, and the research from the perspective of economics mainly focuses on business model, innovation model, enterprise organization, and other fields. First, regarding research on the business model, Arnold et al. [1] divided the business model of the industrial Internet of Things into cloud-based business model, service-oriented business model, and process-oriented business model according to different service objects. Gierej [2] believes that the emergence of the industrial Internet of Things on the one hand heralds the prospect of the development of industrial intelligence and on the other hand marks the transformation of the traditional manufacturing industry to "Outcome Economy," that is, the product and service delivery mode with customer demand as the core. Montori et al. [3] proposed that the industrial Internet of Things has spawned a new business model, which is due to the fundamental change of consumer behavior and the entry of competitors

from outside the manufacturing industry (such as technology enterprises and information technology enterprises). In order to cope with the drastic change of market environment and the increasingly fierce competition, manufacturing enterprises must adjust their original product-centered business model to evolve into the business model of "integration of product and service" [4, 5]. Second, regarding research on the innovation model, Kazuyuki [6] analyzed the relationship between the utilization status, scope, and innovation of big data of Japanese B2B and B2C large enterprises based on questionnaire survey data and pointed out that the application of Internet of Things and big data can on the one hand improve enterprise productivity and on the other hand change enterprise production mode and work content, and the society needs to cope with and adapt to this change. Third, regarding enterprise organization research, Iivari et al. [7] pointed out that the key issues that industrial IoT enterprises need to consider are not only the technical level but also the organizational level so as to achieve the balance and adaptation of enterprise organization, human resources, and industrial IoT technology. This requires the corresponding adjustment of the organizational structure of the enterprise, such as the development and integration of industrial Internet technology, the reconstruction of the core business of the enterprise, and the new mode of production and employee incentive system. Since the beginning of the twenty-first century, academic research on platform enterprises has increased dramatically. The references research path is as follows [8, 9]: since 2000, bilateral market research [10, 11], platform architecture research [9], open innovation research [12], and business ecosystem research have transferred to user innovation research [8, 13], it has formed the research trend of platform strategy under the interdisciplinary background. As soon as the concept of cloud manufacturing was put forward, it became an emerging topic and was warmly welcomed by many scientific researchers. Based on cloud computing, Internet of Things, and other contemporary information technologies, cloud manufacturing extends the concept of "software as a service," an emerging concept in the field of cloud computing, and is introduced into the manufacturing industry.

The research significance of information push technology is mainly reflected in three aspects. First of all, for service providers, a fast and accurate personalized service push system can improve the service quality, provide convenient and high-quality service content for service demanders smoothly, improve their competitiveness in the market, and promote the long-term development of enterprises. Secondly, for the service demand side, the intelligent service push technology saves users' time and energy and improves users' sense of pleasure in experience, and the efficient information push technology can realize the sharing value of information so that the majority of users can quickly find their service needs through this technology. Finally, for manufacturing production form, the development of information push technology accelerates the upgrade of the traditional manufacturing industry and makes manufacturing better integrated into the global economic

development in order to improve the utilization rate of resources, reduce the manufacturing cost, promote the wisdom of the worldwide manufacturing model, and provide strong technical support [14].

From passive search to active screening and elimination, active service push technology has changed greatly the people's lifestyle. With the acceleration of social development, people hope to obtain more convenient and satisfying services through active service push technology. Information push is an active, intelligent, and perceptive information service form. According to the needs of users, the technology combines various advanced information technologies such as big data, artificial intelligence, and other technologies to timely and actively transfer information to the needs of users. This technology is widely used in e-commerce, social networks, information science, user personalized knowledge services, and other fields [15].

Active service push technology combined with automatic system recognition and analysis technology analyzes the potential interest preference of push objects. It helps manufacturing enterprises find potential service objects among a large number of users, realize real-time push operation of personalized service, and improve the accuracy and satisfaction of push service. Active push service is widely used in a web browser and e-commerce platform. For example, Baidu platform analyzes the frequency of browsing and length of stay according to the types of web pages frequently browsed by users and recommends news content that may be of interest to users according to the specific ranking order, thus eliminating the user input keywords or related website.

There are two characteristics of the active service push technology. (1) Active service push refers to the fact that the system automatically obtains the information of the web page browsing by users and uses various advanced data analysis technologies to push the content that users may be interested in. In the past, the demand analysis of services was initiated by users. (2) Cloud platforms are built through service virtualization. It is a gathering place of services and resources, realizing the sharing of services and resources, providing products and services needed by the market efficiently, and delivering them to users timely and accurately. Platform system 24 hours uninterrupted background operation.

Research institutions have done a lot of research on the combination of information push with business processes such as enterprise design and manufacturing. The existing research is mainly reflected in the application of knowledge push in specific fields [16]. In terms of active push of cloud services, Palmisano [17] explored the blank field of online e-retail system and designed and built a recommendation system based on environment awareness. Tosi [18] proposed a method of knowledge push based on user content after identifying the user's interest characteristics. Kobayashi et al. [19] proposed a search engine method based on user interest characteristics in the face of information flooding in the current network environment.

There is a lot of networked manufacturing service information in the cloud manufacturing service platform. This

paper takes the data maintenance disaster caused by big data as the research background, analyzes the historical browsing data of users, and speculates users' preferences, tastes, and needs from the perspective of operators. A machine learning method based on the Markov chain model is proposed to aggregate users with the same interest characteristics, and then a collaborative filtering recommendation algorithm is used to analyze browsing logs, interaction logs, and transaction logs of users with the same interest characteristics. Finally, according to users' ratings, recommend some novel service information that users have not contacted before or may be interested in.

3. Construction of Network Propagation Model

This section introduces the peripheral architecture of the network propagation and how to use machine learning to realize the active push process of service.

3.1. The Peripheral Architecture of Network Propagation. As shown in Figure 1, the peripheral architecture of the push system is composed of four modules, which are data capture, data preprocessing, user clustering analysis with similar interest preferences, and service push module.

The main functions of each module in the push system are as follows.

3.1.1. Data Capture Module. This module provides important user data resources for simulation experiments. The data acquisition of this module mainly includes two processes: (1) web crawler technology to simulate logging into a social manufacturing network to obtain interested topic database information; (2) web crawler technology combined with a website API interface to capture the information of many ordinary users in the platform. After a month of data collection, the web crawler helped obtain personal and browsing information from unauthorized users, eventually collecting 830 users and more than 200,000 text messages. In order to reduce potential errors, this paper divides the data set into 20% as test set and 80% as training set according to the rule of 2 : 8.

3.1.2. Data Preprocessing Module. The module mainly refines and analyzes the obtained text information and converts the obtained text information into numerical information to enable computer recognition. This module mainly includes two processes. (1) Transform the interest topic library information into a standard user interest eigenvalue vector. (2) With the standard user interest eigenvalue vector as the reference, the text information of each user is transformed into the eigenvalue vector as the input parameter of machine learning.

3.1.3. User Clustering Module with Similar Interest. This module mainly gathers a class of users with the same interest and hobbies together through the Markov chain model in machine learning, reduces the scope of user collaborative
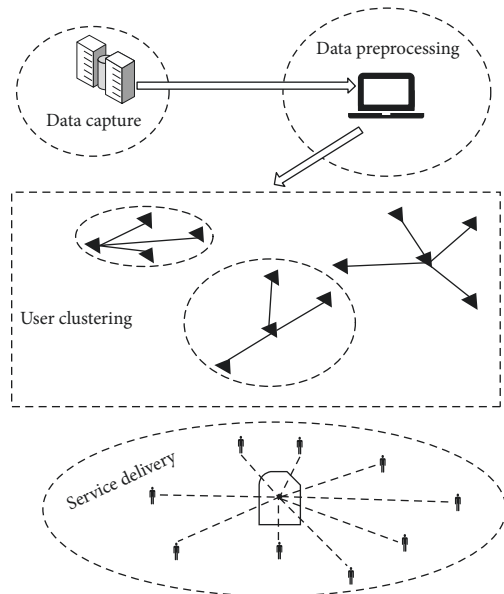
Figure 1: The peripheral framework of the push system.

filtering calculation, and reduces the workload of the working platform. This module takes the user interest eigenvalue vector obtained in the data preprocessing module as the relevant parameters of the multi-Markov model. In the multi-Markov model, the eigenvalue vector of each user can be regarded as a Markov chain, so the set of users becomes the set of Markov chains. In the set of Markov chains, clustering and merging Markov chains with the same or similar characteristics can realize the aggregation of a class of users with the same interests.

3.1.4. Service Push Module. This module is based on user clustering module with similar interests. Through the clustering module, a class of users with the same interest characteristics are gathered together. By analyzing the transaction record information obtained by the data acquisition module, the user-based collaborative filtering recommendation method is used to realize the active push of the service.

3.2. Overall Framework Design of Push System. Section 3.1 outlines the peripheral architecture of the push system, and this section will analyze the main contents of the push system module one by one. Figure 2 describes the main contents of the system in detail. Data acquisition module and data preprocessing module mainly include data acquisition, interest topic database data, and data processing of ordinary users. User clustering module mainly includes transforming user's eigenvalue vector into Markov chain and similar Markov chain clustering operation; service push module mainly includes service evaluation and active push operation of user transaction records. This section will introduce the main contents of user clustering module and service push module in detail.
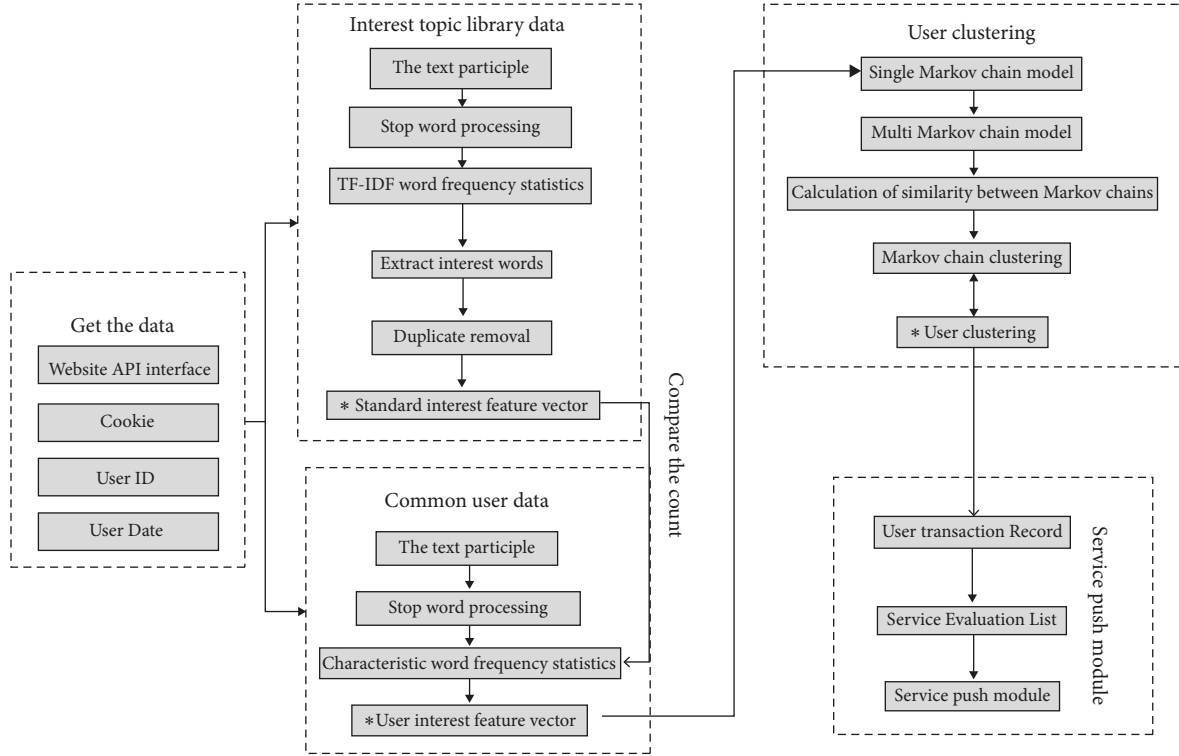
FIGURE 2: Push system internal details.

3.3. Markov Chain Model

3.3.1. Single Markov Chain Model (s-MCM). According to the characteristics of the Markov chain, the interest characteristic value vector of each ordinary user is transformed into the corresponding parameters in the single Markov chain model to predict the user's interest orientation and interest category. However, if you want to use the single Markov chain model to judge the user interest category, you must make the following assumptions about the user browsing process on the website.

Assumption l. (Markov hypothesis) Assuming that the browsing process is a random process, users can be regarded as homogeneous discrete Markov chain. Therefore, a user's browsing record on the website is translated into a sequence of random variables X, which satisfies Markov property.

Under assumption 1, we establish a single Markov chain model (s-MCM).

Definition 1. s-MCM is represented by triple $MC = \langle X, A, \lambda \rangle$ [20]. X represents the discrete random variable, its value range is $\{x_1, \ldots, x_i, \ldots, x_n\}$, and x_i is one state of the model, representing that the user stays in the state of interest. A represents state transition probability matrix, which is also named transition matrix. In the transition matrix, the probability of transition from state x_i to state x_j can be calculated by the formula $P_{ij} = P(X_t = x_j / X_{t-1} = x_i)$, and the transition process can be simplified and written as a state pair (x_i, x_j). λ represents

the initial state distribution, where each term $P_i = P(X_{t=0} = x_i)$. In summary, the transfer matrix A is

$$A = (p_{ij}) = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1j} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2j} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{i1} & P_{i2} & \cdots & P_{ij} & \cdots & P_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & \cdots & P_{nj} & \cdots & P_{nn} \end{bmatrix}, \quad (1)$$

$$R = \begin{matrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{matrix} \begin{bmatrix} 4 & 0 & 0 & 5 & 0 \\ 5 & 5 & 0 & 5 & 4 \\ 0 & 5 & 4 & 0 & 3 \\ 5 & 3 & 5 & 5 & 4 \\ 4 & 0 & 4 & 3 & 0 \end{bmatrix}. \quad (2)$$

At the same time, the initial state distribution is as follows:

$$\lambda = (p_i) = (P_1, P_2, \ldots, P_n). \quad (3)$$

According to the definition of the single Markov chain model, a user's interest eigenvalue sequence X_i is essentially a collection of all pages x_i visited by the user. By counting the number of visits that the user clicks on different pages x_i, the vector of the user's number of visits to different pages is obtained, that is, the interest eigenvalue vector X_i. All the data recorded in the process is denoted d_1. By analogy, the

data recorded by the m th user is d_m, so the data set obtained by all users is $D = \{d_1, d_2, \ldots, d_m\}$.

Thus, the parameters P_{ij} and P_i in the single Markov chain model based on all user data can be estimated by the data set D and the maximum likelihood estimation method; namely,

$$p_{ij} = \frac{s_{ij}}{\sum_{j=1}^n s_{ij}}, \tag{4}$$

$$p_i = \frac{\sum_{j=1}^n s_{ij}}{\sum_{i=1}^n \sum_{j=1}^n s_{ij}}. \tag{5}$$

In (4) and (5), s_{ij} represents the number of occurrences of state pairs (x_i, x_j) in all user text information. When the parameters p_{ij} and p_i are calculated, substituted into (2) and (3), the triple parameters $\langle X, A, \lambda \rangle$ of the single Markov chain model can be obtained based on all user data.

3.3.2. Multi-Markov Chain Model. As described in the summary of Section 3.3.1, it is inaccurate to describe all users' interest characteristics and categories only by establishing a single Markov chain model based on all users' data. In order to improve the accuracy of predicting user interest characteristics, this paper introduces the concept of multi-Markov chain model based on the single Markov chain model, which aims to obtain higher accuracy. The reason why the multi-Markov chain model is more suitable for the modeling of this topic is that the user's interest preference is affected by the user's own manufacturing environment, manufacturing capacity, and other factors. It is an uncertain complex sequence affected by multiple external factors. Therefore, the same category of users with the same or similar interest features is more suitable to be described by the same model; on the contrary, it is more reasonable to describe different types of users with different or not similar interest characteristics by different models. Based on this idea, a multi-Markov chain model based on all user data is established to describe the interest characteristics of all users. The premise is that the user's browsing process needs to meet the following two assumptions.

Assumption 2. (User Classification Hypothesis) Let all users have K categories, and then $C = \{c_1, c_2, \ldots, c_k\}$ denotes the category set of users, and $P(C = c_k)$ denotes the probability that any user belongs to category c_k.

$$\sum_{k=1}^K P(C = c_k) = 1. \tag{6}$$

Assumption 3. (Markov-like chain hypothesis) Let the same category of users have the same or similar interest features, and the sequence of these features is a unique random process. The multi-Markov chain model is defined under Assumptions 1 and 2.

Definition 2. m-MCM is represented by four tuples $\langle X, K, P(C), MC \rangle$. X is a discrete random variable in the range of $\{x_1, \ldots, x_i, \ldots, x_n\}$, and x_i is a state of the model, representing the user interest eigenvalues. K is a preset value in the model, representing the number of user classifications. $C = \{c_1, c_2, \ldots, c_k\}$ denotes the category of users, and the probability distribution of different types of users is represented by the distribution function $P(C)$. Element mc_k represents the Markov chain composed of the same user c_k with the same interest type. If the element mc_k is set together, we obtain the class Markov set $MC = \{mc_1, mc_2, \ldots, mc_k\}$, which represents the set of Markov chains of all categories of users.

In the multi-Markov chain model, the state transition matrix A_k of the class k is

$$A_k = (p_{kij}) = \begin{bmatrix} P_{k11} & P_{k12} & \cdots & P_{k1j} & \cdots & P_{k1n} \\ P_{k21} & P_{k22} & \cdots & P_{k2j} & \cdots & P_{k2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{ki1} & P_{ki2} & \cdots & P_{kij} & \cdots & P_{kin} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{kn1} & P_{kn2} & \cdots & P_{knj} & \cdots & P_{knn} \end{bmatrix}. \tag{7}$$

And the initial distribution of class k is

$$\lambda_k = (p_{ki}) = (P_{k1}, P_{k2}, \ldots, P_{kn}). \tag{8}$$

In (7) and (8), the unknown parameters p_{kij} and p_{ki} can be obtained according to the definition of the multi-Markov chain. Because any user has its own unique interest characteristics in the initial state and can be regarded as a class of users alone, the parameters p_{kij} and p_{ki} can be obtained by a calculation method similar to the single Markov chain; namely,

$$P_{kij} = \frac{S_{kij} + a_{kij}}{\sum_{j=1}^n (S_{kij} + a_{kij})}, \tag{9}$$

$$p_{ki} = \frac{\sum_{j=1}^n (S_{kij} + \alpha_{kij})}{\sum_{i=1}^n \sum_{j=1}^n (S_{kij} + \alpha_{kij})}. \tag{10}$$

In (9) and (10), S_{kij} represents the number of state pairs (x_i, x_j) appearing in the class k user text information. α_{kij} is the superparameter. And the superparameter α_{kij} can be calculated by assuming that all state pairs (x_i, x_j) have the same number of occurrences, that is, Bayes assumption, and then α_{kij} can be obtained by (11) approximation:

$$\alpha_{kij} = \frac{\beta}{n \times n}. \tag{11}$$

In (11), constant β is often replaced by the size n of the problem domain space. In summary, when the parameters p_{kij} and p_{ki} are calculated, substituted into (7) and (8), we obtain the four-tuple parameters of the m-like Markov chain mode $\langle X, K, P(C), MC \rangle$. It is worth declaring that there are m users in the whole process, and the users are divided into m categories; each class Markov chain calculated represents a sequence of user interest eigenvalues.

3.4. Fusion of Interest Feature Vector and Multi-Markov Chain Model. The user's interest eigenvalue vector has been calculated in Section 3. This section will focus on how to use the known interest eigenvalue vector to calculate the parameters in the multi-Markov chain model.

The core idea is summarized as follows. Firstly, $K = m$ is set to generate a quasi-Markov chain, and the practical significance of this step is to regard each user as an independent category, so the sequence of interest characteristic values of m users is correspondingly transformed into m Markov chains. Then, according to the one-to-one correspondence between m users and m Markov chains, the clustering operation of these Markov chains is carried out to realize the clustering of m users. The specific Markov-like chain generation process is as follows.

The initial state is regarded as a multi-Markov chain model under special circumstances; that is, a user's interest eigenvalue sequence is a category ($K = m$), and then the data set $D = \{d_1, d_2, \ldots, d_m\}$ composed of m user's interest eigenvalue is divided into K categories: $D_1, D_2, \ldots, D_k$. Therefore, each user category c_k is essentially a single Markov chain model based on learning data D_k. Therefore, the parameters p_{kij} and p_{ki} can be obtained by (9) and (10), and then the state transition matrix A_k and the initial state λ_k are obtained.

At this point, the interest eigenvalue vector of m users is transformed into a Markov-like chain of m users. Then, the clustering of similar or similar users is actually the clustering of similar or similar user class Markov chain.

3.5. Markov Chain Clustering. Section 3.4 has introduced in detail how the text data of m users are transformed into a Markov chain like m users; that is, the collected user data is transformed into a Markov chain set. The user clustering is actually the clustering of Markov chains, so the similarity evaluation criteria of two Markov chains must be defined in advance.

According to the definition, the dynamic characteristics of the Markov chain are represented by its state transition matrix, so the similarity between Markov chains mc_k and mc_l can be defined by mining the relationship between the two state transition matrices. The definition process of similarity is as follows. Let mc_k and mc_l be any two Markov chains, and their state transition matrices are A_k and A_l, respectively. The i th row p_{ki} parameter in the state transition matrix A_l is $P_{kij} | j = 1, 2, \ldots, n$. And the i th row p_{li} parameter in the state transition matrix A_k is $P_{lij} | j = 1, 2, \ldots, n$.

The parameters of both rows represent the distribution of variables $[t]$ under a given condition $X[t-1] = x_t$, that is, $P(X_t | X_{t-1} = x_t)$, whose approximation can be represented by their cross-entropy. The parameters of the two lines P_{ki} and P_{li} both represent the distribution of the variable $[t]$ under the given $X[t-1] = x_t$ condition, that is, $P(X_t | X_{t-1} = x_t)$, whose approximation can be represented by their cross-entropy:

$$CE(P_{ki}, P_{li}) = \sum_{j=1}^{n} P_{kij} \log \frac{P_{kij}}{P_{lij}}. \tag{12}$$

Therefore, the approximate degree of state transition matrices A_k and A_l can be expressed by the mean of cross-entropy of all row distributions; namely,

$$Similarity(A_k, A_l) = \frac{\sum_{i=1}^{n} CE(P_{ki}, P_{li})}{n}. \tag{13}$$

However, if (4)–(16) is directly used to evaluate the similarity between state transition matrices A_k and A_l; there will be two defects:

(1) First, the cross-entropy (12) is directional, so (13) does not meet the symmetric property; that is, $Similarity(A_k, A_l) \neq Similarity(A_l, A_k)$.

(2) Equation (13) reflects that the more similar the dynamic characteristics of two Markov chains are, the smaller the similarity value of the two Markov chains is, which does not conform to the definition of similarity semantics.

Therefore, to avoid the above two defects, we define the similarity of mc_k and mc_l of two Markov chains as

$$Similarity(mc_k, mc_l) = \frac{2}{Similarity(A_k, A_l) + Similarity(A_l, A_k)}. \tag{14}$$

Note: if the $Similarity(A_k, A_l) + Similarity(A_l, A_k) = 0$, $2/0 = \infty$. The analysis (14) shows that the more similar the dynamic characteristics of two Markov chains are, the greater the similarity value is, which meets the requirement of similarity semantics.

3.6. Multi-Markov Chain Model Clustering Method. Based on the above steps, we give a user clustering method based on the multi-Markov chain model.

(i) User clustering steps:

(ii) Input: user learning data $D = \{d_1, d_2, \ldots, d_m\}$.

(iii) Output: User clustering result.

(iv) Process:

(v) Step l: Use (9) and (10) to map the learning data D into M Markov chains.

(vi) Step 2: Use (11) to (14) to calculate the similarity degree between two Markov chains and arrange them according to the similarity value obtained by calculation from large to small.

(vii) Step 3: According to the preset number of classifications and similarity values, cluster m Markov chains.

(viii) Step 4: According to the corresponding relationship between m Markov chains and m users, output the clustering result.

3.7. Design of Active Service Push Algorithm. The service push system can be vividly described by a black box. The basic data resources of the pushed object are processed by the black box to obtain the push list, and they will receive service

recommendations from the platform at a certain point offline.

In the existing service push system, the system that can achieve accurate push has not emerged because each push method only uses a small part of the system information. To achieve the goal of pushing more accurate, higher quality, and higher satisfaction services for service push objects, this paper adopts multilevel and all-round analysis methods to deal with the information acquired in the system, that is, the cluster-service collaborative filtering and mixed push method. Through the clustering analysis of users in the platform, users with the same interest category are classified into one class. Then, the collaborative filtering recommendation algorithm is used to push services to users with the same interest category. Because the evaluation model of "user-service" for users of the same type is similar, there will not be too many nonevaluation services, which makes the evaluation matrix sparse. This can improve the accuracy of service push. The user clustering analysis method in the mixed push method has been introduced in detail in the previous sections, and the user-based collaborative filtering method will be explained in detail in the following.

3.8. User-Based Collaborative Filtering Recommendation Method.
Through user clustering, a class of nearest neighbor user sets with similar interest characteristics can be obtained. The services in user sets are all closely related to users. Although the massive service types and quantity at the beginning of push have been reduced through clustering operation, service demanders will still be at a loss when facing a series of untargeted push services if these services are sent to service push objects in no order. In order to reduce the time spent by service demanders to understand each service, the nearest neighbor user set is substituted into the "user-service" evaluation model, which aims to generate a ranking list of services that users may like through this model. Finally, the service push object will receive the push information offline and in some way.

3.9. Establishment of "User-Service Evaluation Model".
"User-service evaluation model" is shown as follows. Assuming that there are m users in the nearest neighbor user set and s corresponding service sets, the "user-service" evaluation matrix R is

$$R = \begin{matrix} & \begin{matrix} r_1 & \cdots & r_j & \cdots & r_s \end{matrix} \\ \begin{matrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_m \end{matrix} & \begin{bmatrix} R_{11} & \cdots & R_{1j} & \cdots & R_s \\ \vdots & & & & \vdots \\ R_{i1} & \cdots & R_{ij} & \cdots & R_{is} \\ \vdots & & & & \vdots \\ R_{m1} & \cdots & R_{mj} & \cdots & R_{ms} \end{bmatrix} \end{matrix}. \quad (15)$$

In (15), UI represents the i th user, Rj represents the j th service, and Rij represents the score of the i th user on the j th service.

Rij is a real number ranging from 1 to 5. The higher the value of Rij is, the higher the user satisfaction is. If the user UI does not evaluate the service Rj, the value of Rij is defined as 0. The real-value scoring method is widely spread among service platforms because this scoring mode is more reasonable and detailed, which can greatly reduce the sparsity of "user-service" evaluation matrix R so as to improve the accuracy and satisfaction of service push to the greatest extent.

3.10. Similarity between Computing Services.
When the "user-service" evaluation matrix R is used to calculate the list of push score values of similar services, cosine similarity value and Person similarity value between services can generally be calculated. However, the Person similarity calculation method is more delicate, and the average operation of subtracting the score value from the overall service score value is used to reduce the impact of different evaluation values on the same service with the same evaluation. The calculation results are more reliable and convincing, so the Person similarity measurement is adopted in this paper.

3.11. Person Similarity Measurement Calculation Method.
Assuming that I_{pq} is the set of users who have commented on service r_p and r_q together in the nearest user set and $\overline{R_p}$ and $\overline{R_q}$ are the average ratings of users on service r_p and r_q, respectively, then the similarity $\delta(r_p, r_q)$ is shown in (16):

$$\delta(r_p, r_q) = \frac{\sum_{i \in I_{pq}} (R_{ip} - \overline{R_p})(R_{iq} - \overline{R_q})}{\sqrt{\sum_{i \in I_{pq}} (R_{ip} - \overline{R_p})^2} \sqrt{\sum_{i \in I_{pq}} (R_{iq} - \overline{R_q})^2}}. \quad (16)$$

3.12. The Calculation Example of Service Actively Pushes.
Calculation example 1: assume that users $u_0, u_1, u_2, u_3, u_4, u_5$ are a small group of nearest neighbor users after a clustering, and its user-service evaluation matrix is

$$R = \begin{matrix} & \begin{matrix} r_1 & r_2 & r_3 & r_4 & r_5 \end{matrix} \\ \begin{matrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{matrix} & \begin{bmatrix} 4 & 0 & 0 & 5 & 0 \\ 5 & 5 & 0 & 5 & 4 \\ 0 & 5 & 4 & 0 & 3 \\ 5 & 3 & 5 & 5 & 4 \\ 4 & 0 & 4 & 3 & 0 \end{bmatrix} \end{matrix}. \quad (17)$$

Among them, $\overline{R_1} = 3.67, \overline{R_2} = 2.83, \overline{R_3} = 2.67, \overline{R_4} = 3.0, \overline{R_5} = 1.83$. The users who jointly commented on service 1 and service 2 are u_1, u_2, u_4, so $\delta(r_1, r_2)$ can be calculated by the following formula:

$$\delta(r_1, r_2) = \frac{\sum_{i \in I_{12}} (R_{i1} - \overline{R_1})(R_{i2} - \overline{R_2})}{\sqrt{\sum_{i \in I_{12}} (R_{i1} - \overline{R}_1)^2} \sqrt{\sum_{i \in I_{12}} (R_{i2} - \overline{R}_2)^2}}$$

$$= \frac{(5 - 3.67) \times (5 - 2.83) + (4 - 3.67) \times (4 - 2.83) + (5 - 3.67) \times (3 - 2.83)}{\sqrt{(5 - 3.67)^2 + (4 - 3.67)^2 + (5 - 3.67)^2} \times \sqrt{(5 - 2.83)^2 + (4 - 2.83)^2 + (3 - 2.83)^2}} \tag{18}$$

$$= 0.7413.$$

Furthermore, the calculation process of $\delta(r_1, r_3), \delta(r_1, r_4)$, and $\delta(r_1, r_5)$ is as follows:

$$\delta(r_1, r_3) = \frac{(4 - 3.67)(4 - 2.67) + (5 - 3.67)(5 - 2.67) + (4 - 3.67)(4 - 2.67)}{\sqrt{(4 - 3.67)^2 + (5 - 3.67)^2 + (4 - 3.67)^2} \times \sqrt{(4 - 2.67)^2 + (5 - 2.67)^2 + (4 - 2.67)^2}} = 0.9422,$$

$$\delta(r_1, r_4) = \frac{(4 - 3.67) \times (5 - 3) + (5 - 3.67) \times (5 - 3) + (5 - 3.67) \times (5 - 3) + (4 - 3.67) \times (3 - 3)}{\sqrt{(4 - 3.67)^2 + (5 - 3.67)^2 + (v)^2 + (4 - 3.67)^2} \times \sqrt{(5 - 3)^2 + (5 - 3)^2 + (5 - 3)^2 + 0}} = 0.8908, \tag{19}$$

$$\delta(r_1, r_5) = \frac{(5 - 3.67) \times (4 - 1.83) + (4 - 3.67) \times (3 - 1.83) + (5 - 3.67) \times (4 - 1.83)}{\sqrt{(5 - 3.67)^2 + (4 - 3.67)^2 + (5 - 3.67)^2} \times \sqrt{(4 - 1.83)^2 + (3 - 1.83)^2 + (4 - 1.83)^2}} = 0.9819,$$

where $\delta(r_1, r_5) > \delta(r_1, r_3) > \delta(r_1, r_4)\delta(r_1, r_2)$ is sorted according to the service similarity value. And the active service push list is shown in Table 1.

Therefore, according to the service push data in Table 1, if user u_0 is the active service push object, service 5 and service 3 will be pushed preferentially.

4. Experimental Analysis

4.1. Experimental Environment and Data. Experimental data set is a user data resource, which is obtained through the API interface provided by the official platform of China Machinery Network. In this paper, the web crawler software developed by myself crawls more than ten official users selected initially and extends other user data according to these users, realizing layer-by-layer crawler and deep mining. By the time of simulation analysis, about 830 users had been mined and recorded, who posted more than 200,000 SMS messages. In the face of such a large data set involving all aspects of society, it is crucial to correctly identify valid and invalid data. First, the distribution of the amount of text information published by these users is graphically described as follows.

As shown in Figure 3, the number of SMS messages sent by users is close to 248, and the number of similar users in this range is 478. When fewer than 200 messages were posted, the number of received messages dropped sharply, with the lowest users posting only about six messages. For example, the more detailed the user's information, the more accurate the classification of the user's interest characteristics, and vice versa. Therefore, in subsequent experimental tests, users with less information need to be

filtered to improve the accuracy of user interest feature classification.

4.2. Experimental Evaluation Method. After the implementation of user clustering and service active push operation, this paper adopts the classic mean absolute error (MAE) as the measurement index to evaluate the accuracy of the recommendation system to verify its accuracy. MAE is an effective measure to evaluate the difference between the predicted score and the real score. Its basic idea is to calculate the difference between the predicted score and the real score and absolute value. After the absolute value operation, the evaluation indexes are all positive numbers, which avoids the offset between positive and negative values so as to reflect the deviation of predicted value error from the real score value to the greatest extent. The mathematical formula of MAE is defined as follows:

$$\text{MAE} = \frac{1}{N} \sum |P_{ij} - R_{ij}|. \tag{20}$$

In (17), N represents the sum of the scores predicted by the test, P_{ij} represents the predicted score value of user i for service j, R_{ij} represents the real score of user i for service j. According to (17), the smaller the MAE value is, the smaller the predicted value error is and the higher the accuracy of recommendation is, and vice versa.

4.3. Experimental Analysis

4.3.1. Data Dimension Reduction. Section 3.2 introduces the processing method of interest subject database information

Table 1: Active service push list.

Push sequence sort	Service type	Service evaluation
1	r_5 (service 5)	0.9819
2	r_3 (service 3)	0.9422
3	r_4 (service 4)	0.8908
4	r_2 (service 2)	0.7413

and user text information in detail, and we obtain the interest eigenvalue vector of each user. In Section 4.3.2, the multi-Markov chain model is established to convert the eigenvalue vector of each user's interest into a unique corresponding state transition matrix, whose dimension is 624×624. Because each user has a unique Markov chain state transition matrix and the matrix dimension is 624×624, the clustering calculation will produce repeated data calculation and data dimension disaster, which will greatly affect the calculation efficiency. To sum up, in order to reduce the server load and improve the efficiency of calculation, this paper needs before the user clustering for all users of the Markov chain state transition matrix PCA (Principal Component Analysis) dimensionality reduction. The advantages of PCA dimension reduction are as follows: (1) data dimension reduction can alleviate dimensional disasters; (2) data dimension reduction can minimize information loss while compressing data; (3) the structure of data after dimensionality reduction is easier to understand through visualization. In short, the operation essence of PCA is to map the original data to the low-latitude numerical space through a series of numerical linear transformations under the premise that the characteristics and attributes of the original data are not damaged as much as possible. Figure 4 is the Pareto figure of user state transition matrix.

4.3.2. User Clustering.

As described in Section 4.2, the accuracy of the user clustering effect in the early stage is directly related to the accuracy of user-service active push in the later stage. The clustering of users with the same or similar interest characteristics can increase the accuracy of the later service's active push. Therefore, this section mainly discusses the user clustering effects of the two methods. This paper adopts two methods to realize user clustering: (1) enhanced learning method in machine learning to realize user clustering, namely, multi-Markov chain model clustering method introduced in Section 3 (m-MCM clustering); (2) the unsupervised learning method in machine learning that realizes the clustering of users, namely, the classical K-means clustering method, the implementation of which is provided by Weka software. After user data processing, user clustering results are shown in Table 2.

As can be seen from Table 2, different clustering methods can obtain different clustering results; that is, there will be some deviation in the amount of clustering for each type of user. However, it is difficult to judge which clustering method has a better clustering effect only based on the data in Table 2. Therefore, Figures 5 and 6 are used to show the clustering situation of the K-means method and the
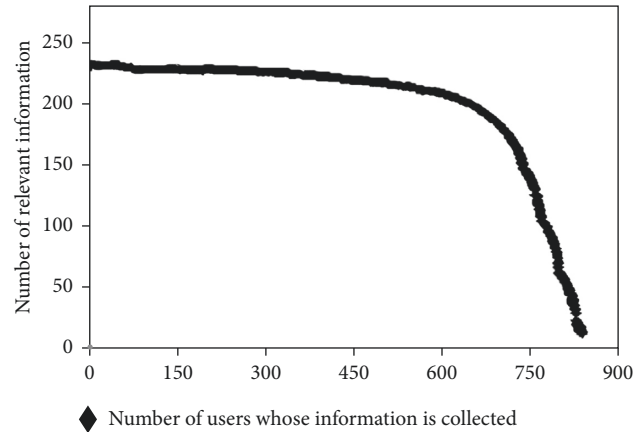
Figure 3: Distribution of the amount of published information of the NTH user.
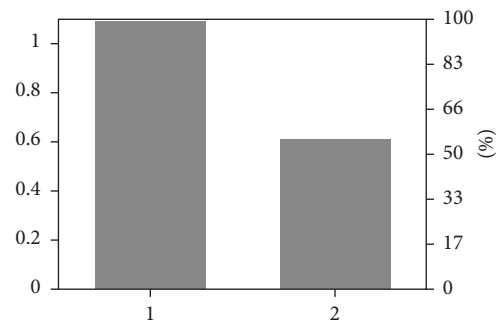
Figure 4: Pareto figure of user state transition matrix.

Table 2: m-MCM and K-means clustering results.

Category	m-MCM	K-means
C_1	56	50
C_2	35	39
C_3	94	35
C_4	75	124
C_5	43	40
C_6	21	26
C_7	68	63
C_8	40	43
C_9	65	69
C_10	31	36
C_11	40	45
C_12	27	12
C_13	38	61
C_14	25	29
C_15	42	37
C_16	31	16
C_17	18	30
C_18	27	24
C_19	19	18
C_20	35	34

m-MCM method for all users, respectively, so as to better distinguish the clustering effect of the two clustering methods. Note: each colored graph in the figure represents a
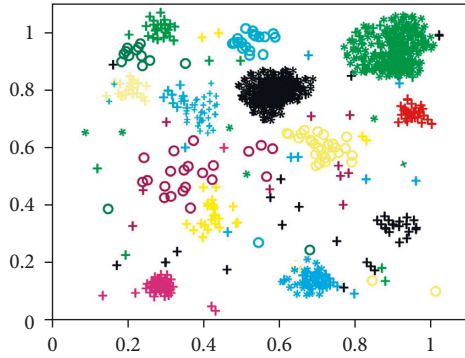
FIGURE 5: The effect of *K*-means user clustering method.
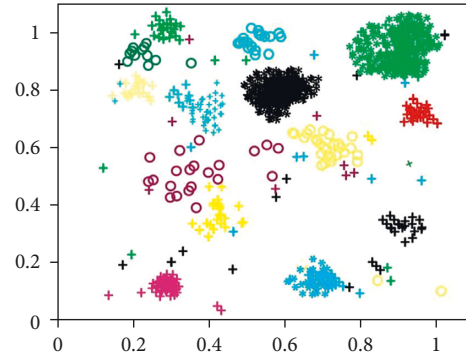
FIGURE 7: The effect picture of *K*-means user clustering method after denoising.
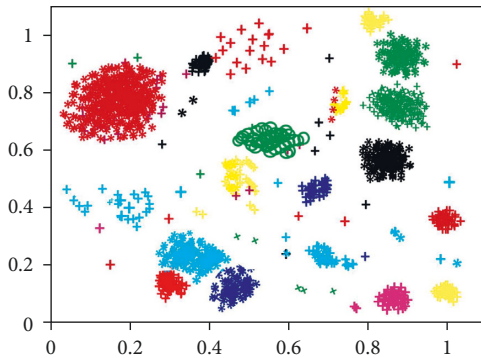
FIGURE 6: The effect of m-MCM means user clustering method.
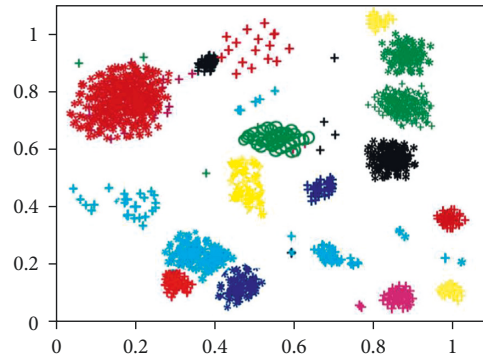
FIGURE 8: The Effect picture of m-MCM user clustering method after denoising.

user, a group of users with the same or similar characteristics of interest.

By observing Figures 5 and 6, it can be found that the scattered noise points in the two images make the clustering effect picture very messy, and it is difficult to clearly distinguish the distribution of clustering users. These noise points are caused by the users with less published information mentioned in Section 3.10. Because the data of these users are less, it is difficult to accurately identify their interests, so they become noise points in the clustering effect figure. Because of these scattered noise points, it is difficult to distinguish the clustering of users accurately. At the same time, noise points will greatly reduce the accuracy of user clustering and affect the clustering effect. Therefore, noise point denoising is also adopted in this paper; that is, users with less data are deleted in advance. The user clustering effect figure after denoising is shown in Figures 7 and 8.

After denoising, Figures 7 and 8 show an obvious improvement in user clustering effect compared with Figures 5 and 6 without denoising. Although there are still noise points in Figures 7 and 8, the clustering of users can be distinguished, so the existence of some noise points is acceptable and within the error range. However, through observation, it can be found that there are significant differences in the user clustering effect in Figures 7 and 8. These differences come from the following: (1) In Figures 7 and 8, the clearance between cluster and the cluster is bigger, and it shows that other users have been spun off the accurate

segmentation. On the contrary, in Figure 7, the gap between clusters is very small, and some clusters are even closely connected, which makes it difficult for us to determine which cluster these points close to the boundary belong to. Therefore, the m-MCM method has a better user clustering effect than the *K*-means method in the gap comparison between clusters and families. (2) As shown in Figure 8, within each family, the distance between points is mostly compact. On the contrary, in Figure 7, the distance between points inside each cluster is too scattered, which makes it difficult for us to distinguish some close clusters and determine whether this cluster is an aggregation of one category cluster or multiple category clusters. Therefore, the m-MCM method has a better user clustering effect than the K-means method in terms of gap comparison within clusters. (3) According to the clustering distribution in Figure 8, we can accurately divide the distribution of 20 types of user interest characteristics. On the contrary, according to the clustering effect in Figure 7, we can only roughly distinguish the user clustering distribution of 18 categories, which is a little deviation from the 20 categories of interest subject database that we set in advance.

In addition, Figure 9 lists the comparison figure of clustering time consuming of *K*-means method and m-MCM method. The figure shows that the clustering time
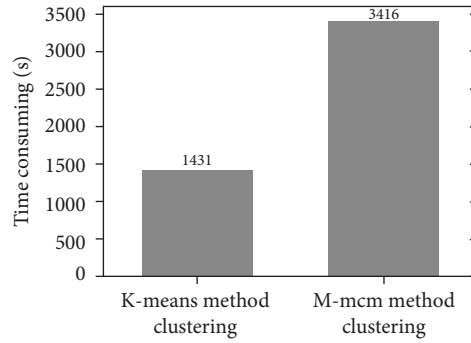
FIGURE 9: Comparison of clustering time between *K*-means method and m-MCM method.

No clustering method

K-means method clustering

M-mcm method clustering
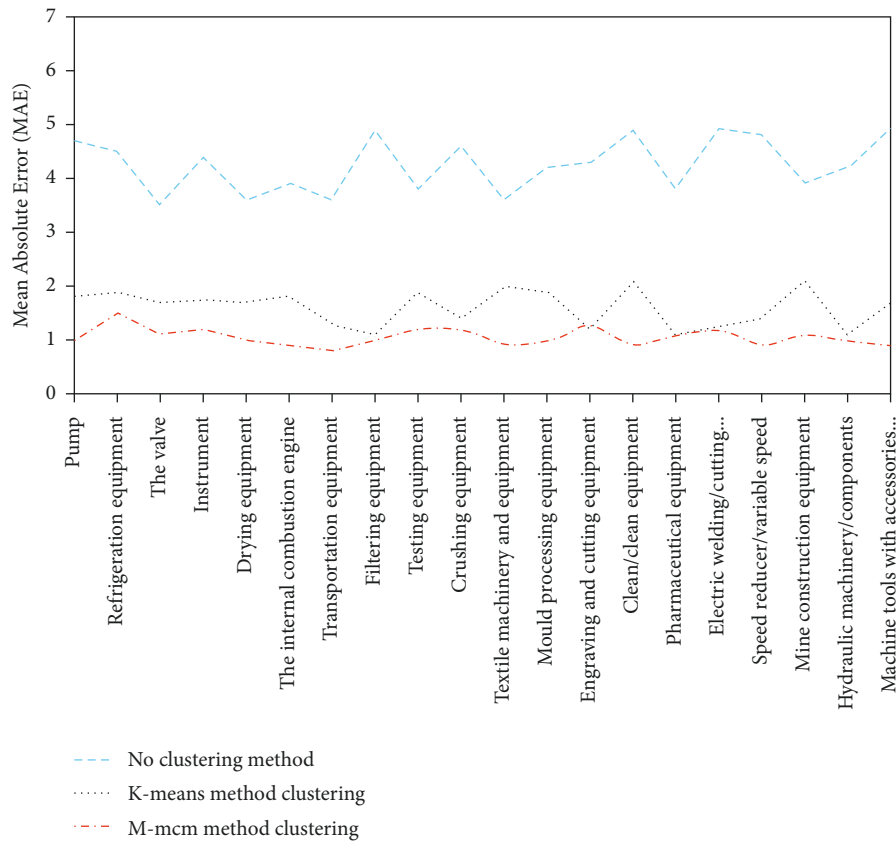
FIGURE 10: MAE under different clustering methods.

of the m-MCM method is 3416 seconds, while that of the *K*-means method is 1431 seconds; that is, the clustering time required by the m-MCM method is about two times that of the *K*-means method. The m-MCM method is not dominant in the time consumed by clustering because the m-MCM method needs to repeatedly calculate the similarity between two Markov chains corresponding to 830 users in clustering. Therefore, although m-MCM clustering has a good clustering effect, it takes a huge amount of time.

To sum up, the m-MCM method is better than the *K*-means method in the user clustering effect. The m-MCM method can relatively accurately determine the interest category to which users belong and accurately define the boundary between classification and class, but it takes more

time. If the m-MCM method is applied to the active push of user service, the accuracy of push can be greatly improved.

4.4. Verification of Collaborative Filtering Recommendation Method Based on User Score Clustering. In order to verify the conjecture that the clustering operation in the early stage has a significant impact on the collaborative recommendation in the later stage, a series of simulation experiments are conducted to compare and evaluate the correctness of the conjecture based on the experimental results. Before the beginning of the experiment, the evaluation method of the experiment, data sources, and preprocessing methods should be explained in advance: (1) The evaluation method

of the experiment is measured by MAE in Section 4.2. The smaller the MAE is, the higher the accuracy of the recommendation is. (2) The experimental evaluation results were measured according to the MAE value of each category. (3) The actual score in the experiment comes from the user's "user-service" evaluation matrix, while the predicted score is the service score calculated by the "user-service" evaluation model. (4) The scoring data of the experiment is about 160,000 service scoring data of 830 users in the social manufacturing platform. (5) The extended service set shall be considered after user clustering. If the original user does not have this service, the score of the extended service shall be set as 0. If the original user has a score for this service score, it will be added to the "user-service" evaluation matrix according to the original service score. As described in the previous chapter, the better and more accurate the user clustering effect in the early stage, the higher the accuracy of user-service push in the later stage. Therefore, three small experiments are designed in this section to verify this assumption: (1) without user clustering, collaborative filtering recommendation is carried out directly for users; (2) cluster users by K-means method, and then make collaborative filtering recommendation; (3) cluster users using the m-MCM method, and then make collaborative filtering recommendation. Experimental evaluation results are shown in Figure 10.

Through the analysis of Figure 10, the following conclusions can be drawn: (1) MAE value obtained by the nonclustering method is much larger than the MAE value obtained by the clustering method; that is, the accuracy of service push can be greatly improved by clustering users first and then actively pushing service to users. (2) The MAE value obtained by clustering with the m-MCM method is smaller than that obtained by clustering with the K-means method. In other words, different methods of clustering users will result in different accuracy of service push. At the same time, the better the clustering effect, the higher the accuracy of service push. (3) In comparison with MAE values of certain categories, the MAE values obtained by K-means clustering are smaller than those obtained by m-MCM clustering, which indicates that K-means clustering is also advantageous in some fields. If different clustering methods can be reasonably used in different fields to actively push users' services, in this way, a more precise and active service push effect can be achieved, which also provides a new research idea for future research. (4) As mentioned above, the user clustering takes a quantity, and the m-MCM method of clustering is not dominant. However, the time spent by M-MCM method is almost twice that of K-means method, but in return, the service push accuracy of M-MCM method is also close to twice that of K-means, and higher accuracy of service push will also consume more time. Therefore, these costs are acceptable in terms of time.

5. Summary and Discussion

The experiment in Section 4.2 implements user clustering and active service push. On this basis, the experiment also compares the clustering effect of the K-means method and the m-MCM method, as well as the accuracy of active service push. The experiment draws the following conclusions: (1) Cluster analysis of users before collaborative filtering can greatly improve the accuracy of active service push. (2) Different user clustering methods lead to different active service push. If the user clustering effect is better, the accuracy of the service push will be higher. Through experimental verification, this paper concludes that the clustering effect of the m-MCM method is better than that of the K-means method, and the accuracy of the clustering effect of the m-MCM method is better than the accuracy of service push after K-means method clustering. (3) In the early stage, spending more time on user clustering can improve the accuracy of the cluster, which is of great help to the later active service push. Since this paper constructed the active service push system for the first time, the experiment in Section 4.2 only preliminarily completed the basic architecture of active service push. In the future, the push system should be further updated and improved in the following aspects: (1) Reduce the clustering analysis time in the early stage of the push system and improve the clustering accuracy without reducing the user clustering effect. (2) A combination of multiple clustering methods is adopted to conduct cluster analysis on users so as to improve the speed and accuracy of cluster analysis and lay a good foundation for later service push. (3) The establishment of a more perfect "user-service" evaluation model can more accurately realize personalized service recommendation for users, thus improving the accuracy of service push by users.

Data Availability

The data can be obtained from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] C. Arnold, D. Kiel, and K. I. Voigt, "Innovative business models for the industrial Internet of Things," *BHM Berg- und Hüttenmännische Monatshefte*, vol. 162, no. 9, pp. 371–381, 2017.

[2] S. Gierej, "The framework of business model in the context of industrial Internet of Things," *Procedia Engineering*, vol. 182, pp. 206–212, 2017.

[3] F. Montori, L. Bedogni, M. DiFelice, and L. Bononi, "Machine-to-Machine Wireless Communication Technologies for the Internet of Things: Taxonomy, Comparison and Open Issues," *Pervasive Mobile Comput*, vol. 50, pp. 56–81, 2018.

[4] V. Krotov, "The Internet of Things and new business opportunities," *Business Horizons*, vol. 60, no. 6, pp. 831–841, 2017.

[5] M. Hartman and B. Halecker, "Management of Innovation in the Industrial Internet of Things," in *Proceedings of the XXVI International Society for Professional Innovation Management (ISPIM) Conference*, Budapest, Hungary, June 2015.

[6] M. Kazuyuki, "Survey of big data use and innovation in Japanese manufacturing firms," *RIETI Policy Discussion Paper Series*, vol. 17, no. 27, pp. 1–24, 2017.

[7] M. M. Iivari, P. Ahokangas, M. Komi, M. Tihnen, and K. Valtanen, "Towards ecosystemic business models in the context of industrial Internet," *Journal of Business Models*, vol. 4, no. 2, pp. 42–59, 2016.

[8] M. A. Cusumano and A. Gawer, "The elements of platform leadership," *MIT Sloan Management Review*, vol. 43, no. 3, pp. 51–58, 2002.

[9] C. Y. Baldwin and C. J. Woodard, *The Architecture of Platforms: A Unified View*, Edward Elgar, Cheltenham, UK and Northampton, US, 2009.

[10] J. C. Rochet and J. Tirole, "Platform competition in two-sided markets," *Journal of the European Economic Association*, vol. 1, no. 4, pp. 990–1029, 2003.

[11] A. Hagiu, "Pricing and commitment by two-sided platforms," *The RAND Journal of Economics*, vol. 37, no. 3, pp. 720–737, 2006.

[12] H. Chesbrough, W. Vanhaverbeke, and J. West, *Open Innovation: Researching a New Paradigm*, Oxford University Press, Oxford, United Kingdom, 2006.

[13] C. Y. Baldwin and E. Von Hippel, "Modeling a paradigm shift: from producer innovation to user and open collaborative innovation," *Organization Science*, vol. 22, no. 6, pp. 1399–1417, 2011.

[14] X. Xu, "From cloud computing to cloud manufacturing," *Robotics and computer-integrated*.

[15] X. V. Wang and X. W. Xu, "DIMP: an interoperable solution for software integration and product data exchange," *Enterprise Information Systems*, vol. 6, no. 3, pp. 291–314, 2012.

[16] A. Pramanik, "Concept of push technology applied in an engineering college library in Kolkata," *Asian Journal of Multidisciplinary Studies*, vol. 2, no. 9, pp. 195–204, 2014.

[17] C. Palmisano, A. Tuzhilin, and M. Gorgoglione, "Using context to improve predictive modeling of customers in personalization applications," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 11, pp. 1535–1549, 2008.

[18] D. Tosi, "An advanced architecture for push services," in *Proceedings of the Fourth International Conference on Web Information Systems Engineering Workshops, 2003*, pp. 193–200, Rome, Italy, December 2003.

[19] I. Kobayashi and M. Saito, "A study on an information recommendation system that provides topical information related to user's inquiry for information retrieval," *New Generation Computing*, vol. 26, no. 1, pp. 39–48, 2007.

[20] J. Borges and M. Levene, "Data mining of user navigation patterns," *Web usage analysis and user profiling*, pp. 92–112, Springer, Berlin Heidelberg, 2000.

Retraction

Retracted: Application of Improved DEA Algorithm in Public Management Problem Classification

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

1. Discrepancies in scope
2. Discrepancies in the description of the research reported
3. Discrepancies between the availability of data and the research described
4. Inappropriate citations
5. Incoherent, meaningless and/or irrelevant content included in the article
6. Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Li, "Application of Improved DEA Algorithm in Public Management Problem Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2568162, 13 pages, 2022.

Research Article

Application of Improved DEA Algorithm in Public Management Problem Classification

Jialin Li [ID]

Vanderbilt University, Nashville, TN 37240, USA

Correspondence should be addressed to Jialin Li; bzxyzb0315@bzu.edu.cn

In order to effectively solve the related problems in the process of public management, this research proposes an improved algorithm technology based on Data Envelopment Analysis (DEA) and classification algorithm. With the help of lever management, this paper further solves and overcomes the problem of DEA algorithm itself and the defect of "relative effectiveness." At the same time, in order to avoid the impact on the input and output indicators, with the help of principal component analysis, taking the performance evaluation of public management departments as the research direction, this paper makes an empirical analysis on the performance evaluation of public management departments. The evaluation results of the index system show that the correlation coefficient between the efficiency value of the initial index system and the efficiency value of optimization 2 is 0.977759, and the correlation coefficient is less than 0.7. The evaluation results are more reasonable than those before the improvement.

1. Introduction

In recent years, with the continuous development of the market economy, the social governance structure has been continuously improved, and the pace of social governance has been gradually accelerated. As an important part of the process of social governance, the importance of public management is becoming more and more obvious. The so-called public management is the scientific management of public affairs, which includes the management of the government, administrative management, urban management, and other aspects. With the continuous deepening and development of public management in China, some major contradictions and problems begin to appear [1]. Such problems as insufficient practical operation means, insufficient governance ability, poor organizational personnel quality, and public management personnel training have seriously affected the improvement of the quality of public management [2–4]. This study is based on these issues, through the integration of Improved DEA algorithm to further improve the public management evaluation system.

On the basis of summarizing the above, foreign scholars elaborated the interaction between the transformation of government functions and the implementation of cultural policies. Xue put forward that only by adopting the overall governance mode, crossing over the administrative settings of each department, and realizing "horizontal" cross departmental cooperation can cultural services be truly implemented; otherwise, it will only be "a piece of empty talk" [5]; Zhao believes that the characteristic of cultural output is that with the growth of the national economy, the wage rate of the government's cultural department develops with other departments. However, due to the labor-intensive output, it is difficult for the productivity to keep pace with the wage growth rate, resulting in an increase in the cost of cultural input, which eventually leads to a doubling of public cultural expenditure [6]; Chen proposed that through the joint investment of the government and local cultural institutions, public cultural services can achieve twice the result with half the effort, and cause the domino effect, bringing more investment to cultural undertakings; joint investment requires both sides to express their own views and reach unity through consultation based on the principle

of mutual cooperation and communication. Such a model is conducive to the diversity and innovation of public cultural development [7]; Ma European cultural identity, which involves ethnic and regional identity in a broader sense, is a model of self-understanding through which individuals, groups, and societies define themselves and their relationships with others. This changing relationship, in turn, will affect the change of self-understanding, that is, the difference in government system and environment will lead to the change of cultural identity [8].

In general, the lack of specialized research on the state of foreign service culture also means that research on the development of China's cultural service needs to be improved renewed in thought and action.

In China, DEA has been applied in many fields. National cotton mills, industryers, and smelters use the DEA method to evaluate the performance of companies and deliver the results to the actual department, which has been praised by stakeholders affected. The National Academic Association approached the DEA to evaluate its effectiveness, which was favored by many. The DEA is a great way to learn about the human resources and potential of coal production in China. Some scholars use the DEA method to determine that "finance is the most important factor affecting the performance of China's high-tech industry and human technology" [9]. He hoped that China's high-tech industry will continue to benefit from the application of science and technology of human resources in the field of performance measurement, innovation baking, and financial management. Some researchers have used the best data shell model to measure the performance of listed companies in the coal industry and estimate the market share and product shortages in the decision-making. They believe that the production efficiency

of most listed coal enterprises is acceptable, but there are great individual differences. Many enterprises still have great room for improvement and put forward suggestions to the administrative departments [10].

2. Improved DEA Algorithm and Classification Algorithm

2.1. Improved DEA Algorithm and Its Application. There are designated decision-making units (DMUs) and design solutions (SU). Their sampling units and components have the same M input and output characteristics, and the input parameters and output parameters are as follows:

$$x_p = \left(x_{1p}, x_{2p}, \ldots, x_{mp}\right)^T, \tag{1}$$

$$y_p = \left(y_{1p}, y_{2p}, \ldots, y_{sp}\right)^T, \tag{2}$$

$$x_j = \left(x_{1j}, x_{2j}, \ldots, \overline{x}_{mj}\right)^T, \tag{3}$$

$$y_j = \left(y_{1j}, y_{2j}, \ldots, \overline{y}_{sj}\right)^T, \tag{4}$$

where equation (1) represents the value of the determination p, which includes the value of the measure; equation (2) represents the value of the p-order, resulting in the measured value; equation (3) represents the value of the jth cell at the input index value; equation (4) represents the value of the jth model unit of the output index value [11].

First, the production possibility set $T(1)$ determined by the sample unit set is expressed as follows under the condition that the axiom system of the production possibility set of the sample unit and the decision-making unit is satisfied:

$$T(1) = \left\{(x, y) \mid x \geq \sum_{j=1}^{\overline{n}} \overline{x}_j \lambda_j, y \leq \sum_{j=1}^{\overline{n}} \overline{y}_j \lambda_j, \delta \sum_{j=1}^{\overline{n}} \lambda_j = \delta, \lambda = \left(\lambda_1, \ldots, \lambda_{\overline{n}}\right) \geq 0\right\}, \tag{5}$$

where $(\overline{x}_j, \overline{y}_j)$ indicates the input and output status in the sample cell set. The input indicator reflects the consumed resources and is also the input data. The output indicator reflects the effectiveness and is the output data. According to the construction method of production frontier, the effective frontier of sample production possibility set $T(1)$ is as follows:

Assumption 1

$$\overline{\omega}, \overline{\mu}, \overline{\mu}_0 \text{ meet } \overline{\omega} > 0, \overline{\mu} > 0. \tag{6}$$

Hyperplane composed simultaneously:

$$L = \left\{(x, y) \mid \omega^{-T} x - \mu^{-T} y - \delta \overline{\mu}_0 = 0\right\}. \tag{7}$$

Meet the following set:

$$T(1) \subset \left\{(x, y) \mid \widehat{\omega}^T x - \widehat{\mu}^T y - \delta \widehat{\mu}_0 \geq 0\right\}, L \cap T(1) \neq \varphi, \tag{8}$$

Then, the set represented by L is the effective surface of the sample possibility set $T(1)$, and the intersection $L \cap T(1)$ is the effective frontier of the sample possibility set $T(1)$.

There are now n decision-making units (DMUs) to be evaluated and n sample units selected. Both DMU and sample units have m input and s output indicators, which are expressed as follows:

$$x_p = \left(x_{1p}, x_{2p}, \ldots, x_{mp}\right)^T > 0, \tag{9}$$

$$y_p = \left(y_{1p}, y_{2p}, \ldots, y_{sp}\right)^T > 0, \tag{10}$$

$$x_j = \left(x_{1j}, x_{2j}, \ldots, \overline{x}_{mj}\right)^T > 0, \tag{11}$$

$$y_j = \left(y_{1j}, y_{2j}, \ldots, \overline{y}_{sj}\right)^T > 0, \tag{12}$$

Here, equation (9) represents the value of the input index and p is the unit value of the decision; formula (10) represents the output value and the unit value determined by p; formula (11) represents the value of the input index and the cell value of model j; formula (12) represents the output value and the j-type cell value [12].

The generalized CCR model for decision unit p is as follows:

$$\left(G - C^2R\right) \begin{cases} \text{Maximize } \mu^T y_p = V(d), \\[1ex] s.t\ \omega^T \overline{x}_j - \mu^T \mathrm{d}\overline{y}_j \geq 0, j = 1, \ldots, \overline{m}, \\[1ex] \omega^T x_p = 1, \\[1ex] \mu, \omega \geq 0, r = 1, \ldots, s; i = 1, \ldots, m, \end{cases} \quad (13)$$

where

$$\omega = (\omega_1, \omega_2, \ldots, \omega_m)^T, \quad (14)$$

$$\mu = (\mu_1, \mu_2, \ldots, \mu_s)^T. \quad (15)$$

Equation (14) represents the input index weight of the selected sample unit, equation (15) represents the output index weight of the selected sample unit, and D is a moving factor greater than 0.

The dual problem model of the generalized model can be expressed as follows:

$$\left(DG - C^2R\right) \begin{cases} \min \theta = D(d), \\[1ex] s.t \sum_{j=1}^{\overline{n}} \overline{x}_j \lambda_j \leq \theta x_p, \\[1ex] \sum_{j=1}^{\overline{n}} d\overline{y}_j \lambda_j \geq y_p, \\[1ex] \lambda_j \geq 0, j = 1, 2, \ldots, \overline{n}. \end{cases} \quad (16)$$

The above formula can prove that $(G - C^2R)$ has an optimal solution.

2.2. Linear Separable Support Vector Classifier. Firstly, Lagrange function is introduced.

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} \alpha_i \left(y_i\left((w \cdot x_i) + b\right) - 1\right), \quad (17)$$

where

$$\alpha = (\alpha_1, \ldots, \alpha_l)^T \in R_+^l. \quad (18)$$

The above formula is Lagrange multiplier. From the definition of duality, we can know that the minimum value of Lagrange function with respect to w, b must be obtained first. By extreme condition,

$$\Delta_b L(w, b, \alpha) = 0, \Delta_w L(w, b, \alpha) = 0. \quad (19)$$

This leads to

$$\sum_{i=1}^{l} y_i \alpha_i = 0. \quad (20)$$

$$w = \sum_{i=1}^{l} y_i \alpha_i x_i. \quad (21)$$

Bring equation (20) into equation (17) and use equation (21) to know that the dual problem of the original optimization problem can be expressed as

$$\max_{\alpha} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j \left(x_i \cdot x_j\right) + \sum_{j=1}^{l} \alpha_j. \quad (22)$$

$$s.t. \sum_{i=1}^{l} y_i \alpha_i = 0\ \alpha_i \geq 0, i = 1, \ldots, l. \quad (23)$$

The algorithm we want can be obtained by converting the objective function into minimization.

3. Performance Evaluation of Public Management Cultural Service Classification Based on Improved DEA Algorithm

3.1. Construction of Index System. We check the impact of indicators on museum performance evaluation according to the commonality of specific indicators. The greater the degree of commonality, the greater the degree of common dependence of the index on the common factor, that is, it is relatively effective to use this common factor to explain the index. Correlation analysis is required before analysis [13]. Generally, the higher the correlation coefficient, the closer the relationship between the two index items. Generally speaking, above 0.7 can be understood as relatively high. By taking the data of 31 provinces in China from 2018 to 2021 as the sample data (124 sample data), after standardized processing, Spss21.0 was run to calculate the common factor, and the following results (Tables 1 to 4) were obtained [14].

The analysis results of gravel diagram are shown in Figure 1.

Sample data passed KMO test. The higher the KMO value, the greater the significance of the changes and the greater the need for analysis. The chart shows that the KMO value is 0.879, indicating that the index is appropriate for analysis. In general, it is shown in Figure 1. Three indicators have been omitted due to differences in three indicators returned this year: cultural monuments, important artifacts, and the number of patents are less than 0.3. Go back to another measurement to get a performance measurement in time. Based on collaborations and releases, 10 museum performance measures were performed as shown in Table 5 [15].

It is scientific and innovative to use factor analysis to select the evaluation indicators of museums. Factor analysis is used to enhance the accuracy and objectivity of the index system [16]. In the actual operation of the museum, there are

TABLE 1: Sample statistics.

	Mean value	Standard deviation	Analysis N
Employees/person	1948.099	1427.7271	124
Senior title	115.639	81.6104	124
Financial appropriation/thousand yuan	250456.422	196399.2388	124
Collection pieces/set	527715.645	397193.3103	124
Number of cultural relics restored this year/set	1088.343	2128.6877	124
Basic display	295.776	571.9660	124
Hold exhibitions	329.343	274.4959	124
Minor visitors/1000	3912.25414523495100	3214.33641453795100	124
Number of visitors/1000	13687.29044512312100	11149.19042516372100	124
Patent	2.487	6.2203	124
Number of completed projects	6.451	8.3224	124
Exhibition room/1000 square meters	173.34782384712390	139.74732354212390	124
Actual floor area/10000 square meters	365.44756383712250	268.74345533717250	124

TABLE 2: KMO and Barrett tests.

Kaiser–Meyer-01kin measure of sampling adequacy		.876
Bartlett's sphericity test	Approximate chi-square	1672.043
	df	78
	Sig.	0.000

TABLE 3: Common factor variance 1 principal component analysis.

	Initial	Extract
Employees/person	1.000	0.654
Senior title	1.000	0.691
Financial appropriation/thousand yuan	1.000	0.666
Collection pieces/set	1.000	0.653
Number of cultural relics restored this year/set	1.000	0.193
Basic display	1.000	0.077
Hold exhibitions	1.000	0.757
Minor visitors/1000	1.000	0.732
Number of visitors/1000	1.000	0.894
Patent	1.000	0.234
Number of completed projects	1.000	0.631
Exhibition room/1000 square meters	1.000	0.821
Actual floor area/10000 square meters	1.000	0.921

many and complex achievement indicators, but too many indicators are easy to form mutual interference between indicators, and it is difficult to find more accurate original data. Factor analysis method is used to eliminate some unnecessary and redundant indicators to achieve the purpose of simplification and optimization.

3.2. Performance Evaluation of Public Cultural Services Based on Methods

3.2.1. Optimization of Input and Output Indexes Based on DEA Method. DEA method is based on known data sets, so the evaluation results directly depend on the selection of input and output indicators. In the chapter, the index system is constructed theoretically and scientifically, but the principle of the method itself on the data is not fully considered. Therefore, the index system should be analyzed from the perspective before performance evaluation. From the perspective of DEA application, the selection of DEA indicators should follow the following basic principles:

Objective measurement. When selecting a measuring device, whether the measurement objective can be achieved, i.e., the choice of the measuring device, the output measurement should be consistent with the functional and objective measurement. Objective assessment requires a variety of materials and tools to describe the whole process [17].

Number of events. Determine the number of units of measurement. Numerous input and output parameters can increase the number of positives, thereby reducing the effectiveness of the DEA method. Measurements should be

kept as simple as possible at the target location. Some researchers believe that all input and output measures should not exceed 1/3 of the scoring sequence [18].

Relevance. Considering the relationship between input indicators and output indicators, the inputs and outputs of DMU are often not isolated [19]. When there is a strong linear correlation between an input indicator and other input indicators, it can be considered that the information of the indicator has been included by other indicators to a large extent, so it can be considered to exclude the input indicator. The same is true for output indicators. This will also lead to more DEA-effective units and affect the evaluation results.

Diversity. Since the core work of the method is "evaluation," there are generally different aspects under a large evaluation objective. The diversity of input and output indicator systems should be considered [20]. Then, on the premise of determining the evaluation objective, multiple input and output indicator systems should be designed, such as from multiple indicators to fewer indicators, to observe which indicators have the greatest impact on the effectiveness value or design similar but different index systems to reflect different aspects of the evaluation.

TABLE 4: Common factor variance principal component analysis.

Ingredients	Initial eigenvalue			Extract sum of squares load		
	Total	Variance%	Cumulative%	Total	Variance%	Cumulative%
1	7.943	61.161	61.161	7.943	61.161	61.161
2	0.993	7.612	68.774			
3	0.978	7.531	76.213			
4	0.743	5.682	81.943			
5	0.621	4.634	86.876			
6	0.454	3.421	90.122			
7	0.412	3.139	93.201			
8	0.271	2.008	95.212			
9	0.258	1.921	97.136			
10	0.191	1.446	98.599			
11	0.131	.994	99.601			
12	0.034	.241	99.846			
13	0.023	.167	100.00			
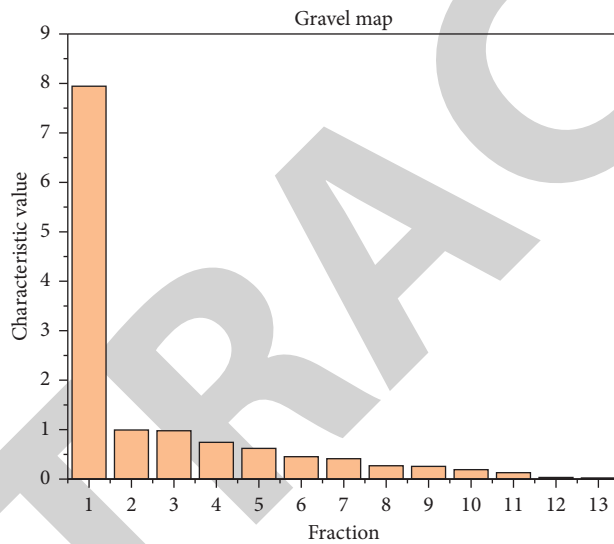
FIGURE 1: Gravel diagram.

TABLE 5: Practical indicators of museum performance evaluation.

Indicator classification	Indicator name
Input indicators	Employees
	Senior title
	Financial appropriation
	Number of collections
	Actual floor area
Output indicators	Number of collections
	Hold exhibitions
	Number of visitors
	Number of completed projects
	Minor visitors

The specific process is as follows: calculate the effective value of each decision-making unit through the practice index system and then calculate the effective value of each decision-making unit through the optimization index system that specifically reflects the evaluation purpose. Finally, the correlation coefficients of the two sets of effective values are calculated. If there is no significant correlation between the effective values calculated by the

two index systems, it can be considered that the optimization index system is not comprehensive. On the contrary, if the correlation is significant, it indicates that the optimized index system can replace the previous index system for performance evaluation analysis [21]. The steps are as follows:

Step 1: determine the practice index system and preliminarily evaluate each decision-making unit; Step 2: optimize the index system based on principles and criteria and conduct preliminary evaluation on each decision-making unit; Step 3: calculate the correlation of the effectiveness values of the two sets of index systems obtained by the method and use the above criteria to judge the comprehensiveness and equivalence of the two sets of index systems so as to form a relatively accurate index system.

3.2.2. Principle of Determining Index Weight

(1) The principle of system optimization. In the evaluation index system, each index has its role and contribution to the system and has its importance to the system. Therefore, the determination of index weight should not only start from a certain index but should consider comprehensively, deal with the relationship between the evaluation indexes, and reasonably allocate their weights. We should follow the principle of system optimization and take the overall optimization as the starting point and goal. Under the guidance of this principle, analyze and compare each evaluation index in the evaluation index system, weigh their respective roles and effects on the whole, and then judge their relative importance. Determine their respective weights, that is, they cannot be evenly distributed, nor can they unilaterally emphasize the optimization of a certain index or a single index while ignoring the development of other aspects. In practice, each indicator should play its due role.

(2) The principle of combining the subjective intention of the evaluator with the objective situation. The weight of the evaluation index reflects the guiding intention and values of the evaluator and the organization. When they feel that a certain indicator is very important and need to highlight its role, they must use a larger weight for each indicator. However, the actual situation is often not completely consistent with people's subjective will. For example, when determining the weight, we should consider the following issues: historical indicators and realistic indicators, recognized by the society and the particularity of the enterprise, and balance between the same industry and the same type of work. Therefore, we must consider the reality at the same time and combine the guiding intention with the reality. As mentioned earlier, economic and social benefits should be considered simultaneously in evaluating the performance of public cultural services.

(3) The principle of combining democracy with centralism. Weight is people's understanding of the importance of evaluation indicators, is the quantification of qualitative judgment, and is often affected by personal subjective factors. Different people have their own views on the same thing, and they are often different, in which there are reasonable elements. Of course, there are also prejudices caused by personal values, abilities, and attitudes. This requires the implementation of the principle of group decision-making, integrating the opinions of relevant personnel to complement each other and form a unified plan. This process has the following advantages: considering problems comprehensively, making the weight distribution more reasonable, and preventing individuals from understanding and dealing with problems unilaterally. It objectively coordinates the contradictions of different opinions among the evaluation parties. The scheme determined through discussion, consultation, and investigation of various specific situations is very persuasive and eliminates many unnecessary disputes in advance. This is a mode of participating in management. During the discussion of the scheme, all parties put forward their own opinions and have further experience and understanding of the evaluation objectives and system objectives. In daily work, they can better work according to the original objectives [22].

The standard values of qualitative indicators are different from those of quantitative indicators, which are easy to collect. They are generally obtained through the following methods:

(1) Expert experience. Expert experience refers to the experience of experts in judging the performance of the public resource supply based on their own experience, combining the political and economic development situation at that time, as well as the economic and social benefits generated by similar projects, units, or departments using similar funds in previous years, and combining certain domestic and foreign experience.

The questionnaire test establishes qualitative standards through public judgment for some indicators related to public satisfaction and expenditure targets to be achieved.

(2) Horizontal comparison. Make a comprehensive comparison of the results achieved by the supply performance of similar public goods at home and abroad. Of course, with the continuous progress of science and technology in human society, new technologies and methods will continue to emerge, and some qualitative indicators that are difficult to quantify will be gradually quantified or accurately grasped.

3.2.3. Performance Evaluation Method System of Public Cultural Services. When engaging in any work, we must understand the methods. As the resource input of public cultural services depends on the government and financial input, the output mainly depends on the public's awareness and satisfaction with culture. As a result, the competition mechanism in the whole field is not strong, and public cultural services are difficult to quantify. In particular, the public cultural services studied in this paper are

difficult to measure by market means. In addition, the authenticity and comparability of data in the field of cultural public services and the difficulty in quantifying most of the achievements of cultural services make it difficult to measure the performance of the public sector. Despite this, people still try to apply various advanced methods to the performance evaluation of public cultural services. The main methods are as follows:

(1) Performance evaluation table method. The performance evaluation table method, also known as the scoring table method, should be said to be an early method. It mainly uses the relevant performance factors formulated by the government to evaluate the performance of the work (such as the number of annual visitors to the museum and the number of books lent out). Compare and score the work performance with the relevant factors one by one and then get the overall results of the work performance. According to the results, it is classified into four levels: excellent, good, qualified, and unqualified.

(2) Management by objectives. It is the most typical result-oriented performance evaluation method. At that time, the objective management method was widely used. The evaluation object of the objective management method was mainly the work performance of employees, that is, focusing on the completion of objectives. Through this method, employees could be promoted to work towards objectives to a certain extent.

(3) Key performance indicator method. The theoretical basis of the process of measuring performance is based on the "February 8 principle" stated by Italian scientist Pareto. In the process of establishing the business value, it is believed that 80% of the work of each department and staff do 20% of the simple behavior, and this 20% capture will understand all. Along with these concepts, key indicators such as feedback and benefits of the organization's internal processes are identified, modeled, accounted for, evaluated, and implemented, and goals of the industry have exploded. This way, it is possible to clearly identify the key roles of responsibility to the director of the company and to develop standards of performance for the employees in the department.

(4) Data envelopment analysis method. Data envelopment analysis (DEA) is developed on the basis of the theory of "relative efficiency evaluation." Its core idea is to estimate the frontier of effective production by observing a group of values about input and output. It is applicable to the effectiveness evaluation of multiple decision-making units with the same type of multiple inputs and outputs. DEA has obvious advantages in efficiency evaluation: first, in the process of evaluation, only input indicators and output indicators need to be determined, without constructing specific production functions or models; secondly, it is suitable for complex objective evaluation with multiple inputs and outputs; then, this method does not require the dimension of the indicators. As long as a specific indicator of the decision-making unit uses the same dimension, the original data can be directly analyzed. The above advantages of DEA method can just meet the requirements of efficiency evaluation of public cultural services. First of all, the government's human, financial, and material inputs in various branches of public cultural services are eventually transformed into various specific public goods and services. Generally, we can only determine which outputs are related to inputs. As for the transformation relationship behind it, it is difficult to quantify and measure in reality; secondly, public cultural services cover a wide range, which belongs to a typical multi-input and multi-output situation.; Moreover, it is inconvenient or impossible to measure the input and output of public cultural service supply in terms of money or some other specific dimension in many cases, and DEA can just avoid this point. Therefore, the DEA method will have a broad application prospect in evaluating the efficiency of government supply of public goods. This paper is also based on this discussion.

3.2.4. Classification Performance Evaluation of Public Cultural Services Based on C^2R Model. First of all, according to the important position of museums, public libraries, and mass art centers in public cultural services, the practice index system has been established according to the fifth chapter. Then, according to the optimization idea of input and output indicators based on DEA method, the practice indicator system is optimized to form a new indicator system, and the performance structure model diagram of public cultural services is constructed as shown in Figure 2.

Finally, two sets of index systems are input through model C^2R for operation to form the corresponding performance evaluation result analysis. The flow chart of public cultural service classification performance evaluation based on C^2R model is shown in Figure 3.

4. Performance Evaluation of Public Cultural Service Classification Based on C^2R Model

4.1. Determination of Input and Output Indicators. The indicator system constructed according to the above is shown in Table 6.

According to the optimization of input and output indicators of the above DEA method, since the employees' indicators already include the indicators of senior employees, the senior employees are eliminated from the perspective of simplification. Similarly, the actual use of the housing area index already includes the exhibition room index, so the exhibition room is excluded. In the output indicators, the number of visits by minors has been reflected in the number of visitors, so the number of minor visits is excluded. The index system after comprehensive optimization is shown in Table 7.

C^2R model is used to calculate the data of the two groups of index systems, and the efficiency value is obtained as shown in Table 8.
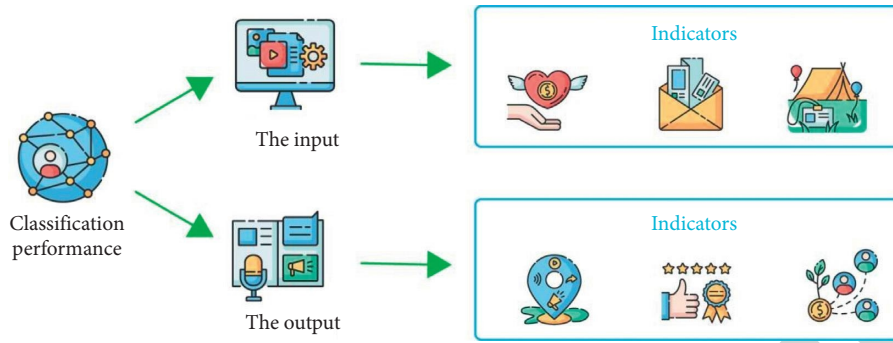
Figure 2: Performance structure model of public cultural services.
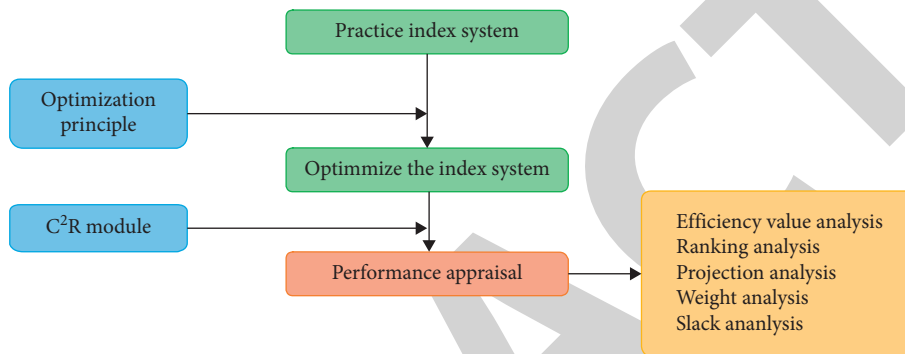
Figure 3: Flow chart of public cultural service classification performance evaluation.

Table 6: Museum performance evaluation index system.

Indicator classification	Indicator name
Input indicators	Employees
	Senior title
	Financial appropriation
	Exhibition room
	Actual floor area
Output indicators	Number of collections
	Hold exhibitions
	Number of visitors
	Number of completed projects
	Minor visitors

Table 7: Optimization indicators for museum performance evaluation.

Indicator classification	Indicator name
Input indicators	Employees
	Financial appropriation
	Actual floor area
Output indicators	Hold exhibitions
	Number of visitors
	Number of completed projects

It cannot be found from the table that the optimized index system has a great impact on some regions, in which the effective value of Beijing is reduced from relatively effective 1 to 0.48, which indicates that the index items subject to subdivision have a great impact on the performance value [23]. The correlation coefficient obtained from the correlation test of the effective values corresponding to the two groups of index systems is 0.644816.

TABLE 8: Cross-value comparison of two index systems.

NO.	DMU	Score	Score-optimize1
1	Beijing	1	0.480356868
2	Tianjin	1	0.537980302
3	Hebei province	0.839556	0.758941004
4	Shanxi province	0.810403	0.552091004
5	Inner Mongolia autonomous region	0.669882	0.568362204
6	Liaoning province	0.486929	0.423907129
7	Jilin province	1	1
8	Heilongjiang province	1	1
9	Shanghai	1	0.532312132
10	Jiangsu province	1	0.973383461
11	Zhejiang province	0.828156	0.799626211
12	Anhui province	1	1
13	Fujian province	0.98733	0.985527088
14	Jiangxi province	0.76535	0.697399767
15	Shandong province	0.908386	0.881542336
16	Henan province	0.921388	0.921406465
17	Hubei province	1	0.834135771
18	Hunan province	1	1
19	Guangdong province	0.92253	0.898512197
20	Guangxi Zhuang autonomous region	0.734522	0.670123232
21	Hainan	1	1
22	Chongqing city	0.966883	0.895422931
23	Sichuan province	1	0.889086523
24	Guizhou province	1	1
25	Yunnan province	1	1
26	Tibet autonomous region	1	1
27	Shaanxi province	0.65188	0.485806762
28	Gansu province	1	0.592953371
29	Qinghai province	1	0.987601029
30	Ningxia Hui autonomous region	0.568312	0.561783341
31	Xinjiang Uygur autonomous region	0.795216	0.703376021

We believe that if the correlation coefficient is greater than 0.7, then our optimized index system cannot replace the original index system. Therefore, it is necessary to further optimize the practice performance evaluation indicators, find out which indicator system has a greater impact, and conduct correlation analysis on the original indicator system as shown in Table 9.

The removed senior staff and exhibition rooms are highly correlated, and the correlation value between the number of collections and other indicators is mostly below 0.7. Therefore, the optimization indicators are re-optimized as shown in Table 10.

C^2R model is used to calculate the data of the two groups of indicator systems, and the effective values obtained are shown in Table 11. The correlation coefficient obtained from the correlation test of the effective values corresponding to the two groups of index systems is 0.903433. We believe that if the correlation coefficient is greater than 0.7, then our secondary optimized index system cannot replace the original index system. Through the optimization of indicators, the indicators after the secondary optimization of museum performance evaluation can relatively comprehensively represent the

indicators of museum performance evaluation. The following analysis is also based on the indicators after the secondary optimization.

4.2. Performance Evaluation Analysis. In the previous section, the efficiency values of provinces are calculated according to the evaluation secondary optimization indicators. After processing, see Figure 4.

The overall service efficiency of national museums will be moderate in 2020. The efficiency level of the national evaluation is 0.870559, and the overall efficiency is moderate. This has something to do with the country's great investment in public cultural undertakings in recent years. However, there are some regions with low efficiency and relatively poor efficiency in Xinjiang and Ningxia. The projection analysis is shown in Table 12, and the investment needs to be adjusted. Through the above index adjustment, we can also find that the number of collections is a very important index item for museums. In the future construction and development of museums, museums should devote more energy to how to enrich the number of collections in museums.

TABLE 9: Correlation analysis of museum index system.

	Employees (person)	Senior title	Financial appropriation (thousand yuan)	Exhibition room (thousand square meters)	Actual floor area (ten thousand square meters)	Number of collections (piece/set)	Hold exhibitions	Number of visitors (thousand)	Number of completed projects	Minor visitors (1000)
Employees (person)	1.00	0.80	0.77	0.87	0.91	0.67	0.76	0.91	0.73	0.78
Senior title	0.81	1.00	0.71	0.84	0.87	0.72	0.82	0.77	0.65	0.66
Financial appropriation (thousand yuan)	0.79	0.72	1.00	0.71	0.81	0.78	0.65	0.67	0.71	0.51
Exhibition room (thousand square meters)	0.85	0.83	0.67	1.00	0.96	0.67	0.92	0.93	0.67	0.84
Actual floor area (ten thousand square meters)	0.90	0.87	0.80	0.98	1.00	0.75	0.92	0.95	0.73	0.82
Number of collections (piece/set)	0.67	0.71	0.78	0.65	0.74	1.00	0.58	0.63	0.61	0.51
Hold exhibitions	0.77	0.82	0.65	0.89	0.89	0.61	1.00	0.87	0.73	0.81
Number of visitors (1000)	0.88	0.77	0.71	0.91	0.92	0.64	0.87	1.00	0.73	0.91
Number of completed projects	0.71	0.67	0.72	0.67	0.77	0.64	0.73	0.74	1.00	0.56
Minor visitors (1000)	0.78	0.63	0.51	0.83	0.81	0.49	0.81	0.94	0.54	1.00

TABLE 10: Secondary optimization indicators for museum performance evaluation.

Indicator classification	Indicator name
Input indicators	Employees Financial appropriation Actual floor area
Output indicators	Hold exhibitions Number of visitors Number of completed projects Number of collections

TABLE 11: Efficiency values of two indicator systems.

NO.	DMU	Score	Score-optimization1
1	Beijing	1	1
2	Tianjin	1	1
3	Hebei province	0.839556	0.758941004
4	Shanxi province	0.810403	0.702091004
5	Inner Mongolia autonomous region	0.669882	0.668362204
6	Liaoning province	0.486929	0.483907129
7	Jilin province	1	1
8	Heilongjiang province	1	1
9	Shanghai	1	1
10	Jiangsu province	1	1
11	Zhejiang province	0.828156	0.829626211
12	Anhui province	1	1
13	Fujian province	0.98733	0.985527088
14	Jiangxi province	0.76535	0.697399767
15	Shandong province	0.908386	0.901542336
16	Henan province	0.921388	0.921406465
17	Hubei province	1	1
18	Hunan province	1	1
19	Guangdong province	0.92253	0.898512197
20	Guangxi Zhuang autonomous region	0.734522	0.730123232
21	Hainan	1	1
22	Chongqing city	0.966883	0.955422931
23	Sichuan province	1	0.909086523
24	Guizhou province	1	1
25	Yunnan province	1	1
26	Tibet autonomous region	1	1
27	Shaanxi province	0.65188	0.602806762
28	Gansu province	1	0.641953371
29	Qinghai province	1	1
30	Ningxia Hui autonomous region	0.568312	0.561783341
31	Xinjiang Uygur autonomous region	0.795516	0.703376061

The regional characteristics of museum performance evaluation are obvious. Among the 31 provinces, autonomous regions, and municipalities directly under the central government in the mainland of China, 14 regions' museum performance is efficient compared with other regions, including Beijing, Tianjin, Shanghai, Hubei, and Anhui. Liaoning, Ningxia Hui autonomous region, Shaanxi, and Gansu have relatively low-efficiency values. The results show obvious regional characteristics.

The low number of collections is the main reason for the low efficiency in some areas, which has been fully reflected in the process of index optimization. Therefore, to improve the efficiency of museums, on the one hand, it is necessary to increase the number of collections, exhibitions, and other measures, on the other hand, it is necessary to make reasonable allocation in terms of employees, financial allocation, and the actual use of housing area.
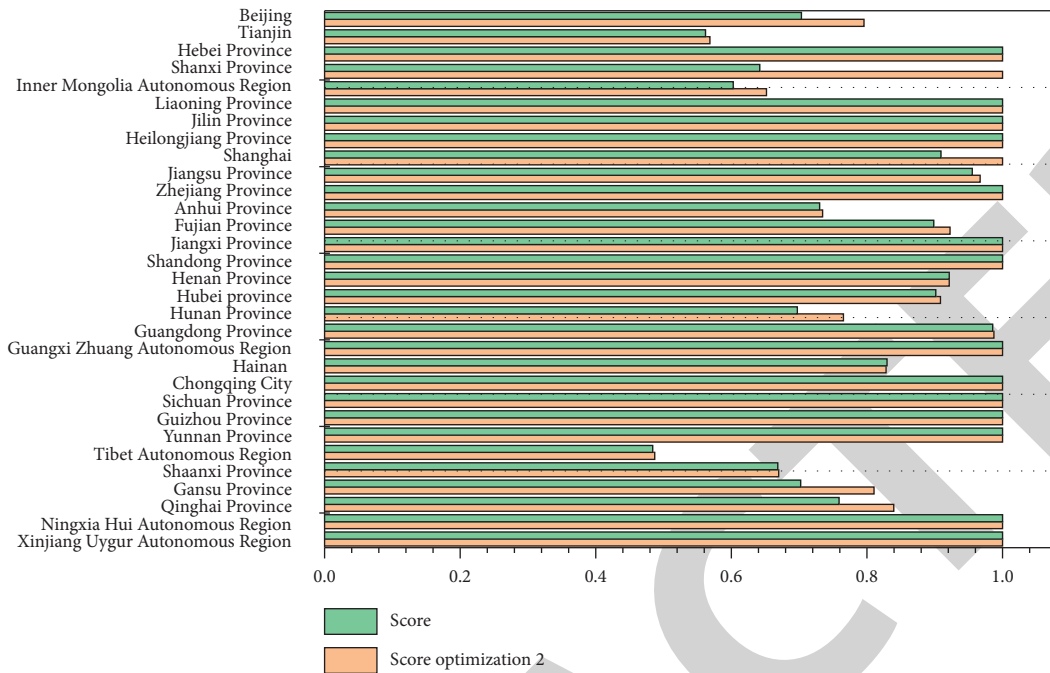
Figure 4: Line chart of museum performance evaluation.

Table 12: Projection analysis of museum performance evaluation.

	Ningxia Hui autonomous region		0.5617833		
	Employees (person)	208	117.41343	−91.587372	−43.81%
	Financial appropriation (thousand yuan)	50246	28227.801	−22019.123	−43.81%
30	Housing area (10000 square meters)	66.17	36.417263	−29.761824	−44.98%
	Hold exhibitions	43	43	0	0.00%
	Number of visitors (thousand)	840.10	840.10	0	0.00%
	Completed projects	0	0.6316811	0.6316811	999.90%
	Xinjiang Uygur autonomous region		0.7023741		
	Employees (person)	853	599.27271	−252.73361	−29.61%
	Financial appropriation (thousand yuan)	126345	88875.084	−37479.921	−29.61%
31	Housing area (10000 square meters)	205.61	117.38304	−87.237841	−42.64%
	Hold exhibitions	183	182	0	0.00%
	Number of visitors (thousand)	6073.74	6073.74	0	0.00%
	Completed projects	1	6.6904798	5.6905199	569.05%

5. Conclusion

Based on the improved DEA algorithm, this research puts forward the theoretical framework of public cultural service performance evaluation system. The theoretical framework of performance evaluation system is the key to the government supply of public cultural services. The construction of the public cultural service performance evaluation index system needs to have a clear target, be good at using scientific analysis methods, scientifically measure the validity and reliability of the index system, and design the public cultural service performance evaluation index system according to the objectives, performance dimensions, performance orientation, and other ideas of performance evaluation. C^2R model and superefficiency model in DEA algorithm are innovatively and comprehensively applied to the performance evaluation of government supply of public cultural services. DEA method has many advantages. It determines the attribution of the index system by fitting the path coefficients of each factor in the model through objective data and automatically generates weights. It is suitable for analyzing the performance evaluation of public cultural services. On the basis of collecting a large number of statistical sample data, taking into account the large differences between different types of public cultural services, this paper makes an empirical analysis on the performance appraisal of public management departments with the help of principal component analysis and taking the performance appraisal of public management departments as the research direction. The evaluation results of the index system show that the correlation coefficient between the efficiency value of the initial index system and the efficiency value of optimization 2 is 0.977759, and the correlation coefficient is less than 0.7. The evaluation results are more reasonable than those before

Research Article

Cultural Heritage Resource Development and Industrial Transformation Resource Value Assessment Based on BP Neural Network

Xinyu Liu [iD],[1] **Yujie Li,**[2] **Zihao Zhang,**[3] **and Qianzheng Wang**[4]

[1]*School of Performance and Cultural Industries, University of Leeds, Leeds LS2 9JT, Leeds, UK*
[2]*College of Science, Xi'an University of Architecture and Technology, Xi'an 710000, Shaanxi, China*
[3]*School of Computing, University of Leeds, Leeds LS2 9JT, Leeds, UK*
[4]*School of Business, University of Leeds, Leeds LS2 9JT, Leeds, UK*

Correspondence should be addressed to Xinyu Liu; zyx223@stu.xjtu.edu.cn

Mining and utilizing cultural heritage resources and creating and developing creative and cultural industries have become the priority direction of economic development, setting off a wave of cultural heritage resource development, and industrial transformation. Which cultural heritage resources can have the high value of industrial transformation has become one of the research topics that have attracted much attention. In view of this problem, research is of great significance to the development of cultural heritage resources and the field of industrial transformation. With the in-depth research on resource development and industrial transformation, the research on artificial neural network (ANN) in cultural industry transformation is gradually carried out, and its performance advantages are of great significance to solve the problem of value evaluation. This paper aims to study the application of the value assessment method based on the BP neural network (BPNN) in the development of cultural heritage resources and industrial transformation. Through the analysis and research of BPNN and cultural heritage resource development and industrial transformation, it can be applied to the construction of resource value assessment methods to solve the problem of improving the value level of cultural heritage resource development and industry. This paper analyzes BPNN, cultural heritage, and value evaluation methods, conducts experimental analysis on the performance of the algorithm, and uses related theoretical formulas to explain. The results show that the evaluation method has passed the random consistency test, and the results are valid. The obtained index popularity weight value is 0.134, the economic benefit weight value is 0.093, and the resource correlation degree weight value is 0.074, which can be used as the key criteria for resource value evaluation. The classification of resources through resource value assessment can provide theoretical support for the development and industrialization of cultural heritage and can meet the needs of improving the value and quality of cultural heritage development and industrial transformation, and the development level and satisfaction have been greatly improved.

1. Introduction

Under the upsurge of cultural resources and heritage protection, developers have carried out development and industrial transformation of traditional cultural heritage. In the past, the development of most of the cultural heritage was too blind, and the value generated by the transformation industry was uneven, and it could not meet the increasing demands of people in terms of economic development and artistic aesthetics. Resource value assessment is an evaluation method to assist decision-making to quantify the value of the assessment object, which can solve the problem of comparing the value of different types of resources. Due to its advantages in comparison, it has been applied to various fields to successfully solve various decision-making problems. It quantifies the value of several different types of resources according to the set indicators according to the purpose of evaluation and judges the value of a certain

resource by sorting the calculated values to assist decision-making. In the development of cultural heritage, its content exists in various forms, and how to identify the development value of a large number of cultural heritage resources has far-reaching significance for its development and industrial transformation. ANN is an operation model, which is composed of a large number of nodes connected to each other. However, ANN has a better effect on the value evaluation problem to be solved and has less restriction, so its application range is very wide. In recent years, scholars have used ANN for cultural resource evaluation, but the application and research of BPNN in this area are relatively few. Therefore, it is of great significance to apply BPNN in this paper to solve the research on resource value evaluation in the development of cultural heritage resources and industrial transformation.

At present, with the continuous advancement of the development of cultural heritage, more and more scholars have explored the development and industrial transformation of cultural heritage resources. Among them, in order to solve the problems of insufficient resource integration and serious destruction of cultural heritage, Pang et al. used the geographic information system platform to establish a basic database of historical heritage [1]. Rouhani introduced the protection and development of cultural heritage in peace-building and economic development and put forward his own views on the path of protection of cultural heritage resources in the future [2]. Yi et al. focused on the latest progress of China's national cultural information resource sharing project to understand people's expectations and requirements for cultural heritage development projects [3]. In order to better develop cultural heritage resources, Heitman et al. studied an extensive and centralized repository of cultural heritage information, providing opportunities for research and facilitating the cocreation of a cultural heritage information exchange community [4]. In order to develop smart cultural tourism services, Ruzic identified and sorted cultural heritage resources that appeared on social media and marked them [5]. However, these are not good for the development of the transformation industry of cultural heritage resources.

BPNN can be used in the value assessment of cultural heritage resources, and it has a good effect on the processing speed of the assessment data and the accuracy of the assessment results. Among them, cultural heritage resources can arouse heated discussions on social platforms. Ding and Tian proposed an algorithm called PRABP to predict the number of retweets in cultural heritage [6]. In order to analyze the tourism volume of cultural tourism resources, Li et al. proposed an effective model based on the Baidu index to predict tourism volume [7]. Zhang et al. proposed a method based on the improved artificial fish swarm algorithm neural network algorithm to improve the prediction accuracy of the audience's preferences for cultural heritage [8]. In order to improve the efficiency of resource value evaluation, Zhang et al. developed a simulator using BPNN to predict the exploitability of cultural heritage resources [9]. These methods improve the efficiency of resource value evaluation to a certain extent, but the accuracy is not high.

In order to solve the problem of the above-mentioned low level of cultural heritage resource development and industrial transformation, this paper uses BPNN to analyze the development of cultural heritage resources and evaluates various indicators to find high-value cultural heritage resources. The innovation of this paper is using BPNN and cultural heritage development and industrial transformation to analyze how BPNN, cultural heritage resources, and value assessment methods play a role in the research on cultural heritage resource development and industrial transformation resource value assessment based on BP neural network. This paper expounds the proposed resource evaluation method. Through the evaluation, this paper finds that this method can improve the accuracy of cultural heritage resource value evaluation and improve the level of industrial transformation.

2. Method of Cultural Heritage Resource Development and Industrial Transformation Resource Value Assessment

The 21st century is an era of cultural and economic integration. The cultural industry and modern technology have become the driving force of economic growth, but the quality of the emerging cultural heritage development projects is uneven, and there are many problems, such as the low actual industrial value of the project and resource damage caused by excessive development. Therefore, it is very important to study the development of cultural heritage resources and the evaluation of the value of industrial transformation resources [10]. The cultural heritage resources are shown in Figure 1: Figure 1(a) is the Fujian Tulou, and Figure 1(b) is the mural of Huashan Mountain.

Through the investigation, it is found that the current research on cultural heritage resource development and industrial transformation resource value assessment based on BP neural network is not complete, and most of the research focuses on the theoretical analysis of cultural resource development strategies or phenomena, so this paper proposes a research on the value assessment of cultural heritage resources using BP neural network [11, 12]. This paper analyzes the related methods of BPNN, cultural heritage resource development, and resource value evaluation and proposes an evaluation method for cultural heritage development and industrialization resource value. Through experimental analysis, it is found that cultural resource value assessment can promote the development of cultural heritage and the better development of industrialization.

This paper mainly introduces the research background of cultural heritage resource development and industrial transformation resource value assessment based on the BP neural network, draws out the problems to be solved to illustrate the purpose and significance of this paper, then makes a general analysis of the research status of cultural heritage resources development and industrialization and the application of BPNN, and explains the content and innovation of this paper. At the same time, the organization structure and method of this paper are described, and the

(a)   (b)

Figure 1: Cultural heritage resources: (a) Fujian Tulou; (b) Huashan Murals.

related methods of BPNN, cultural heritage resource development, and industrialization and resource value assessment are analyzed and described; then, the data source of this paper is explained in detail; after arranging the data, this paper analyzes the development trend of heritage resource development and industrial transformation, the preference of cultural heritage types, the satisfaction of former cultural heritage development projects, and the results of the comparison of various levels within the evaluation index system and different cultural heritage resource projects and draws conclusions; finally, the full text is summarized.

2.1. BPNN Method.

BPNN is an inverse operation of error, which optimizes and repairs the weights of each neuron by sending out information until the independent error of each neuron is obtained and the expected result is obtained. The algorithm has two main processes, one is forward propagation of data, and the other is reverse propagation. When the sample data enter the test program, the final result is sent to the output layer by computing each hidden neuron. The error between the result and the expected result cannot be ignored, and the program enters the antimissile step. The process of this process is mainly to amortize the errors of the output layer to the front through hidden neurons to provide data. On the basis of adjustment and modification, obtaining data according to the output result, and run the program for many times to finally obtain the desired output result. Such resource value calculation will take the influence of each factor into account more comprehensively, as the weights between the various factors need to be modified several times to obtain valid procedural results [13].

The neural network is developed under the inspiration of the neurons of the human brain, and it forms its own processing method by systematically and abstractly processing the data. A neuron is the smallest unit of structure in BPNN [14]. Its structure is shown in Figure 2.

The input of data can be expressed as a stimulus to neurons, and the output of data can be expressed as simple responses of neurons [15]. The relationship between them is shown in the following formulas:
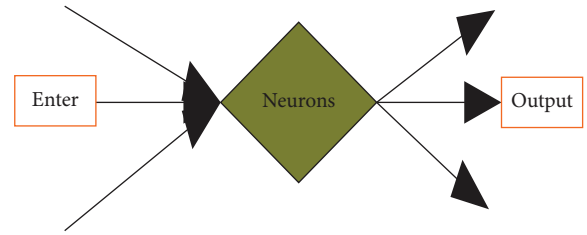
Figure 2: Neuron structure.

$$R_j \sum_{j=1}^{n} q_{ij} z_j - \alpha_j, \tag{1}$$

$$o_i = h(R_i). \tag{2}$$

Among them, z_j is the input variable; α_j is the threshold; q_{ij} is the weight of adjacent neurons; $h(R_i)$ is the transfer function.

BPNN means a multilayer feed-forward network that only propagates the signal from the back to the front [16]. BPNN supports N-to-N multipoint input and multipoint output mapping with a very high degree of nonlinearity. Its basic structure is shown in Figure 3.

As shown in Figure 3, the neural network training is mainly carried out in the data processing process. Through the forward feedback of the network, the model automatically processes the operation to reduce the error and change the weight between the model layer and each node. Stop training if the threshold feedback errors for the layers and the model's control variables are within the desired accuracy [17, 18]. In practical applications, various training methods can be used to shorten the time of model training, optimize the effect of model training, and further improve the convergence speed of model training. The most commonly used method to improve model training efficiency is gradient descent. It is a first-order optimization algorithm, which iteratively searches for the specified step distance point in the opposite direction of the gradient corresponding to the current point on the function to find the minimum value.
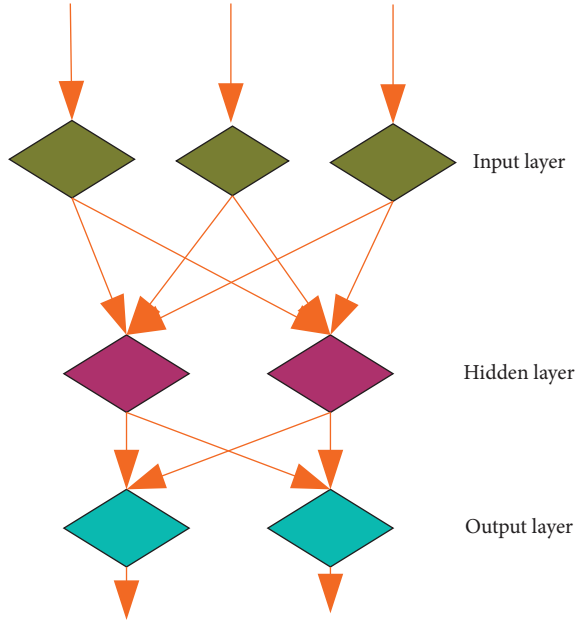
FIGURE 3: Network structure.

The neurons of BPNN learn general linear functions by default as basic functions. Its expression is shown in the following formula:

$$u = \sum_{j=1}^{n} q_j z_j - \alpha = z' q - \alpha. \tag{3}$$

In order to meet the requirements of increasing partial differential equations and monotonicity for specific intervals, the Sigmund function can be used as the output function of the BPNN. The expression is shown in the following formula:

$$h(z) = \left(1 + e^{-\beta z}\right)^{-1}. \tag{4}$$

Then, the derivative function expression is obtained as shown in the following formula:

$$h'(x) = h(z)(1 - h(z)). \tag{5}$$

The forward propagation of data information is analyzed as follows.

The data input from the first-layer network will be processed by the first-layer and second-layer network connection functions, and the process is shown in the following formulas:

$$nt_j = \sum_{i=0}^{n} r_{ij} z_i \, (j = 1, 2 \cdots m), \tag{6}$$

$$u_j = h(nt_j) \, (j = 1, 2 \cdots m). \tag{7}$$

The processed information is then output through the connectivity between the second- and third-layer networks to complete the simple forward propagation of data information. The process is shown in the following formulas:

$$nt_k = \sum_{j=0}^{m} r_{ik} u_j \, (k = 1, 2 \cdots l), \tag{8}$$

$$o_k = h(nt_k) \, (k = 1, 2 \cdots l). \tag{9}$$

The error back-propagation processing of data information is analyzed as follows.

Assuming the sample is obtained as shown in the following formula:

$$Y^d = \frac{1}{2} \sum_{K=1}^{L} \left(a_k^d - o_k^d\right)^2. \tag{10}$$

Substituting it into the calculation formula of the second to third layers to obtain the following formula:

$$Y = \frac{1}{2} \sum_{K=1}^{L} \left(a_k - h\left(\sum_{j=0}^{m} q_{jk} u_j\right)\right)^2. \tag{11}$$

Among them, q_{ij} is the weight of each layer, that is, the error, which can be changed by adjusting the weight.

In general, it is hoped that the weight value should be proportional to the negative gradient of the error, and this is due to the continuous reduction of the error in the learning process, so it is obtained as shown in the following formulas:

$$\Delta q_{jk} = -\delta \frac{\varepsilon Y}{\varepsilon q_{jk}} \, (j = 0, 1, 2, \cdots, m; k = 1, 2, \cdots, l), \tag{12}$$

$$\Delta r_{ij} = -\delta \frac{\varepsilon Y}{\varepsilon r_{ij}} \, (i = 0, 1, 2, \cdots, n; j = 1, 2, \cdots, m). \tag{13}$$

Among them, δ is the learning rate.

Therefore, for the output layer, the formula (14) is obtained as

$$\Delta q_{jk} = \delta \varphi_{jk} u_j. \tag{14}$$

Among them, φ_{jk} is the error signal. The error signal is shown in the following formula:

$$\varphi_{jk} = (a_k - o_k) o_k (1 - o_k). \tag{15}$$

Then, the hidden layer is obtained as shown in the following formula:

$$\Delta r_{ij} = \delta \sum_{k=1}^{l} \varphi_{jk} q_{jk} u_j (1 - u_j) Z_i = \delta \varphi_{jk} z_i. \tag{16}$$

The error signal is then corrected as shown in the following formula:

$$\varphi_{ij} = \sum_{k=1}^{b} \varphi_{jk} q_{jk} u_j (1 - u_j). \tag{17}$$

Through the above formula, the error adjustment of the second and third layers of the sample can be calculated as shown in the following formulas:

Cultural and radio
facilities

Ancient artifacts

Cultural Resources

Intangible cultural
heritage

Celebrities

FIGURE 4: Types of cultural resources.

$$q_{jk}(b+1) = q_{jk}(b) + \delta \sum_{d=1}^{D} \varphi_{ij} u_j, \tag{18}$$

$$r_{ij}(b+1) = r_{jk}(b) + \varnothing \sum_{d=1}^{D} \varphi_{ij} u_j. \tag{19}$$

Since the construction of the BPNN model, many fields have been using it for different degrees of research and development, and the applicability of BPNN in resource value evaluation has been enhanced. When using BPNN to deal with complex resource information, regardless of economic, legal, or technical aspects, no matter how large the dependence between the data is, it can make the model predicted value closer to the function of the real value. So, this paper applies it to the value assessment of cultural heritage resources.

2.2. Cultural Heritage Development and Industrialization Resource Value Assessment. Cultural resources are the first link in the evolution of cultural forms in the production process of cultural industries and are the basis for creativity. Therefore, evaluating the value of local cultural resources is the first step in the development of local cultural industries. Based on the inherent spiritual and material dual attributes of cultural resources, evaluating and measuring their value has a dual meaning. There are many kinds of cultural resources, and their classification is shown in Figure 4.

As shown in Figure 4, the correct classification of cultural resources is an important prerequisite for the development of cultural resources statistics, and research and cultural industry development strategies. According to the definition of cultural resources, avoid using too empty concepts such as material cultural resources, spiritual cultural resources, and intangible cultural resources.

To evaluate the value of cultural resources, we must first clarify the principles of evaluation. Evaluation should strictly adhere to the principle of objectivity, avoid subjective preference guessing, and adopt one-sided methods when evaluating the value of cultural resources and should combine the mainstream value orientation of the historical development of the cultural industry and use scientific methods. Cultural resources should be measured from a realistic perspective, so as to reflect the real value of cultural resources; evaluation should adhere to local principles, and an important task of cultural resource evaluation is to investigate the heritability of dominant factors in regional cultural contexts. Therefore, it is necessary to pay attention to the discovery of core value resources and the thorough discovery of value when evaluating; the evaluation should be based on qualitative and quantitative principles, cultural resource evaluation is based on the weighted design problem commonly used in systems engineering technology. "Analytic Hierarchy Process" can solve multipurpose system evaluation and decision-making problems caused by multipurpose, inclusive, uncertain, complex, and other characteristics. The overall principle should be followed in the assessment. Cultural resources share their own unique cultural

ecological environment. In the industrialization and development of cultural resources, the influence of these cultural and ecological factors needs to be considered. Therefore, assessing the value of cultural resources is not a simple scoring of individual resources but requires looking at it from a holistic point of view.

For the method of resource value evaluation, the first step is to determine the evaluation index of cultural resources and form a set of intuitive evaluation system of cultural resources. Combined with BPNN data analysis, a set of "cultural resources value evaluation index system" is established, and finally, a basic evaluation model is established through the analytic hierarchy process, and this method is used to count and evaluate cultural resources. Through BPNN, the indicators of each layer of the evaluation model are compared and analyzed, and the weight of each indicator is obtained. The weight comparison adopts a top-down approach, determining the weight of upper-level indicators according to the characteristics of cultural resources in industrial development, comparing the weights of lower-level indicators, and refining the top-level indicators to obtain the bottom-level indicators and overall importance. After weighting the indicators of the evaluation model, comprehensively evaluate and analyze the cultural resources of the objects selected in the above-mentioned cultural resources statistics.

3. Data Sources for Cultural Heritage Resource Development and Industrial Transformation Resource Value Assessment

This paper conducts an online questionnaire survey on the types of cultural heritage favored by netizens and their satisfaction with the current industrial transformation, and a total of 348 valid questionnaires are obtained. The specific data results of the survey content are shown in Table 1.

Among them, the content information contained in the questionnaire includes the respondents' gender, age, favorite type of cultural heritage, and their satisfaction with the current cultural heritage resource development project. For the satisfaction evaluation part of the questionnaire, 4 options were set, namely very dissatisfied, general, relatively satisfied, and very satisfied, corresponding to numbers 1, 2, 3, and 4, respectively. According to these data, the needs of cultural heritage resource development and industrial transformation resource value assessment can be analyzed.

At the same time, by taking 10 different types of cultural resource development projects as examples, this paper evaluates the resource value of them and classifies them according to the score of the evaluation value. The cultural resource development project data used for the experiment were obtained through Internet search, and the detailed information data are shown in Table 2.

The content of the table includes the label, project name, province, and time of cultural resource development projects used for value assessment.

Table 1: Partial data of the content of the questionnaire.

Project	Netizen 1	Netizen 2	Netizen 3
Gender	Male	Female	Male
Age	19	41	27
Type of preference	Cultural relics	Ruins	Immaterial
Satisfaction level	1	3	2

4. Results and Discussion of Cultural Heritage Resource Development and Industrial Transformation Resource Value Assessment

4.1. Resource Value Assessment Needs. This paper analyzes the development trend of cultural heritage resource development and industrial transformation from 2012 to 2021 through the data of cultural resource development projects collected by data mining; and through the distribution of online questionnaires, this paper collects the satisfaction evaluation of netizens according to the current development of cultural heritage resources and industrial transformation. After sorting and analyzing the data, the results are obtained. The development trend of cultural heritage resource development and transformation is shown in Figure 5.

Figure 5 shows the proportion of cultural heritage resource development projects from 2012 to 2021. It can be seen that the overall trend is increasing year by year. Figure 5(a) shows that the overall number of projects in the development of cultural heritage resources increased from 2011 to 2016, but the growth rate was slow; Figure 5(b) sees an explosive growth in the number of projects for cultural heritage resource development from 2017 to 2021, with a faster growth rate. This is because people gradually realize the value of the development of cultural resources and heritage and advocate their development and industrialization, setting off a wave of upsurge. It can be seen that in the huge number of cultural heritage development projects, the identification of its development value is particularly important for the effective use of limited resources.

Through the questionnaire survey, netizens were collected about their preferences for cultural heritage types and their satisfaction with the current cultural heritage development projects. The specific results after data collection and arrangement are shown in Figure 6.

Figure 6 shows the results of an online questionnaire survey on the types of cultural heritage favored by netizens and their satisfaction with the current industrial transformation. From Figure 6(a), we can see the distribution of netizens' preferences for the type of resources. Netizens who like cultural relics accounted for 24.4%, those who like ancient buildings accounted for 11.4%, and those who liked sites accounted for 16.4%, and netizens who liked nonmaterial types accounted for 47.8%. It can be seen that the intangible cultural heritage resources are the most popular among the many types of cultural heritage resource development. Figure 6(b) shows the degree of satisfaction of netizens with the current cultural resource development project, among which 23.4% of netizens said they were very dissatisfied, 44.6% said they were average, 21.4% said they were relatively satisfied, and

TABLE 2: Ten different types of cultural resource development projects.

Serial number	Project name	Province	Time
X1	Huangzhong embroidery	Qinghai	2008.6.7
X2	Wudu mountain opera	Gansu	2012.8.7
X3	Emei martial arts	Sichuan	2009.4.12
X4	Painting of Tibetan colored sand mandala	Tibet	2013.2.4
X5	Rice cultivation customs in Wannian County	Jiangxi	2020.4.6
X6	Gansu gaolan shilan ancient pear garden	Gansu	2016.3.14
X7	Traditional incense-making techniques	Guangdong	2019.6.7
X8	Evenki clothing	Inner Mongolia	2014.3.6
X9	Lightning protection project of the fourth detachment of the new fourth army in Shu	Shandong	2020.8.6
X10	Ancient Jujube Garden in Jiaxian County, Shaanxi Province	Shaanxi	2018.5.14
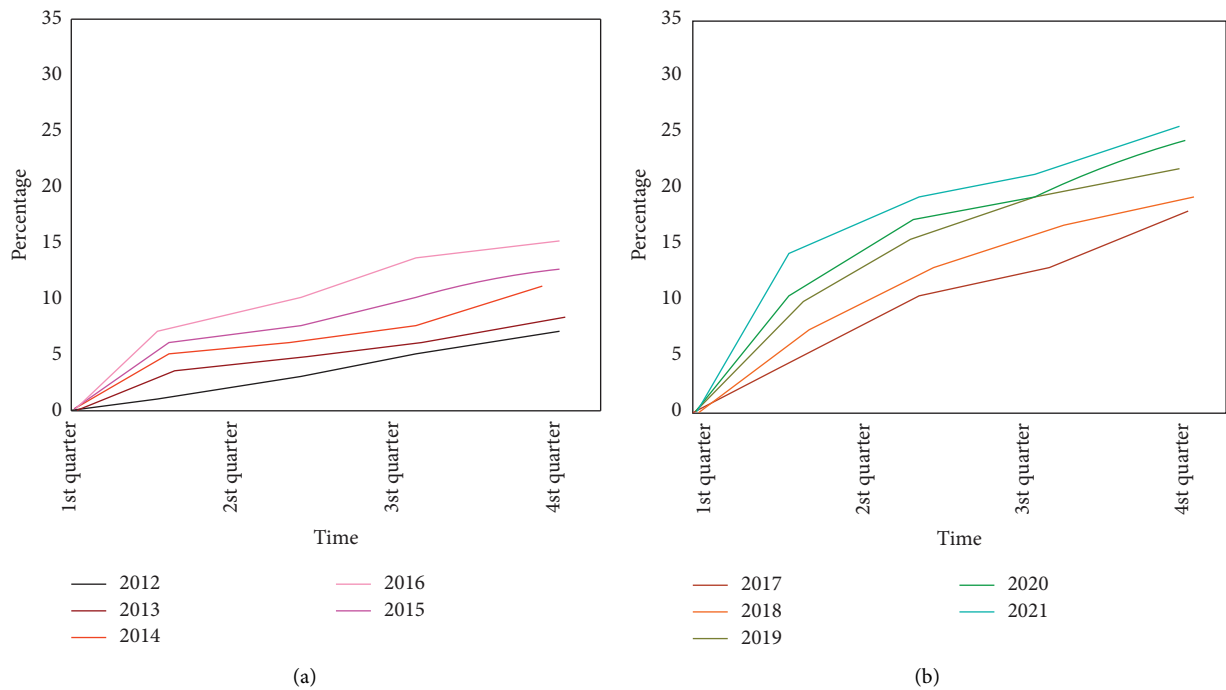
FIGURE 5: Development trend of cultural heritage resource development projects from 2012 to 2021. (a) 2012–2016, (b) 2017–2021.

10.6% said they were very satisfied, the proportion of netizens who expressed dissatisfaction reached 68%. It can be seen that the quality of the current numerous cultural resource development projects is uneven, and netizens are generally not satisfied with them. Therefore, it is necessary to evaluate the value of the cultural resources to be developed before the project development to improve the quality of the project and improve satisfaction.

4.2. Comparison of Various Levels within the Evaluation Index System. This paper calculates the weights of indicators at various levels of cultural resource value evaluation, organizes and analyzes the data, and finally obtains the results shown in Figure 7.

Figure 7 shows the weights of different levels of indicators for the evaluation of cultural resource development value. Figure 7(a) shows that in the criterion layer, the weight of historical heritage value is 0.35, the weight of economic development value is 0.3, the weight of scientific,

educational, and cultural value is 0.25, and the weight of artistic aesthetic value is 0.1. It can be seen that the historical heritage value of the resource has the greatest impact on the evaluation score, and the artistic aesthetic value has the least impact on the result. Figure 7(b) shows the top 6 weight indicators that have a great influence in the project layer, among which the popularity is 0.134, the education and scientific examination is 0.116, the economic benefit is 0.093, the brand effect is 0.083, the ancient degree is 0.098, and the resource correlation degree is 0.074. It can be seen that focusing on the above six indicators in the development and industrialization of cultural resources can improve the actual quality of the project.

4.3. Comparison of Different Cultural Heritage Resource Projects. This paper evaluates the industrialization development value of a total of 10 cultural heritage resource development projects from X1–X10, and the total evaluation results are shown in Figure 8.

(a)

(b)
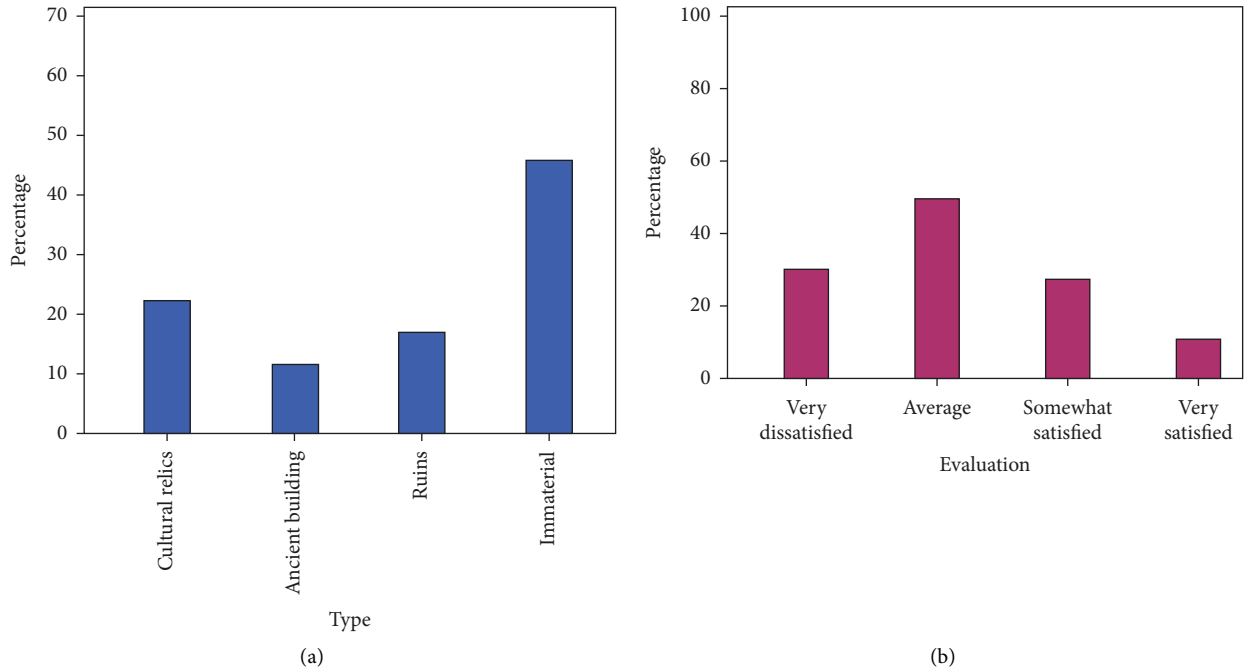
FIGURE 6: Questionnaire survey results. (a) Type preference distribution. (b) Satisfaction survey.

(a)

(b)
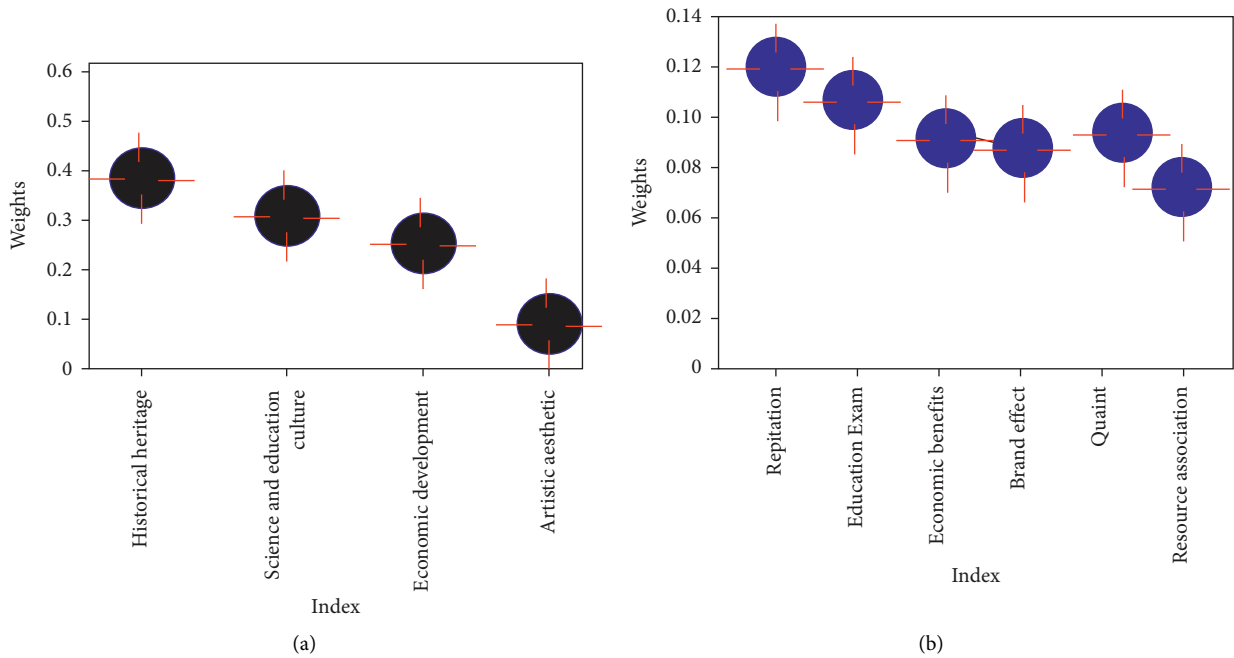
FIGURE 7: Indicator weights at each level. (a) Criterion layer. (b) Project layer.

As shown in Figure 8, on the whole, it can be seen that the value evaluation results of different types of cultural heritage resource development projects are still quite different. Most of the items were scored between 50 and 70 points, with the highest overall score for the assessment of item X7 being 73.2 points and the lowest overall score for the assessment of item X5 being 42.3 points.

According to the results of evaluation and scoring, they can be divided into three categories: semidevelopment

category (1), pending development category (2), and key protection category (3), as shown in Table 3.

Category 1 in Table 3 is a semideveloped category, and the score range for this category is >70 points. This category is a cultural resource project that is in the development stage and has already achieved certain results. It includes Huangzhong embroidery project, Wudu Mountain Opera project, and traditional incense technology, which are high-industrial value projects. Category 2 is the category to be
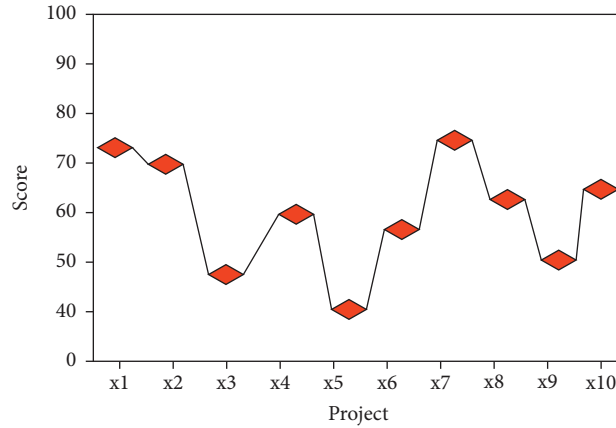
Figure 8: Value assessment results.

Table 3: Classification of the value assessment of the industrialization development of cultural heritage resource projects.

Serial number	Project name	Category
X1	Huangzhong embroidery	1
X2	Wudu Mountain Opera	1
X3	Emei martial arts	3
X4	Painting of Tibetan colored sand mandala	2
X5	Rice cultivation customs in Wannian County	3
X6	Gansu Gaolan Shilan Ancient Pear Garden	2
X7	Traditional incense making techniques	1
X8	Evenki clothing	2
X9	Lightning protection project of the fourth detachment of the new fourth army in Shu	2
X10	Ancient Jujube Garden in Jiaxian County, Shaanxi Province	2

developed, and the score range for this category is 50–70 points. This category is a cultural resource project that can be developed but has not been industrialized for some reasons, or has just started. It includes the Tibetan Caisha Mandala Painting Project, the Ewenki Clothing Project, the New Fourth Army Shu Four Detachment Lightning Protection Project Site Project, the Shaanxi Jiaxian Ancient Jujube Garden Project, and the Gansu Gaolan Shilan Ancient Pear Garden Project, which are projects of medium industrialization value. Category 3 is the protection category, and the score range for this category is <50 points. Compared with the first two types, this category is a cultural resource with lower economic benefits and more serious damage. It includes Emei Wushu Project and Wannian County Rice Farming Custom Project, which are projects of low industrial value but in urgent need of protection.

5. Conclusions

Cultural heritage resources are the treasures of human material and spiritual civilization, and people's requirements for cultural resource development projects are getting higher and higher. The development of cultural heritage resources industrialization is inseparable from the contribution of BPNN. BPNN has been applied in the value evaluation of cultural resources industry development because of its powerful data processing advantages. It can be seen from the comprehensive experimental test that the accurate evaluation of resource value can better promote and improve the quality level of cultural resource industrialization. Through the analysis of the needs of cultural resource value evaluation, it is found that due to the continuous development of the cultural economy industry, a large number of industrialization projects for the development of cultural resources have emerged, so it is particularly important to screen the resource value of these projects. At the same time, in the face of a large number of projects, the quality is uneven. This paper investigates the satisfaction of netizens with the project through a questionnaire survey. Netizens are not satisfied with the current cultural resource development, 68% of netizens expressed dissatisfaction. It can be seen that evaluating the cultural resource value of development projects to improve project quality is also very important to improve satisfaction. By comparing the internal levels of the evaluation index system, it is found that the weight of the influence index on the value of the popularity is 0.134, the education and scientific examination is 0.116, the economic benefit is 0.093, the brand effect is 0.083, the ancient degree is 0.098, and the resource correlation degree is 0.074, and these 6 indicators will be the focus of evaluating the value of resources. By comparing different cultural heritage resource projects, 10 different heritage resource projects are divided into 3 categories according to their scores. Among them, Huangzhong embroidery project,

Wudu Mountain Opera project, and traditional incense technology belong to the development category, with a score range of >70 minute. This type of project has a high degree of resource relationship, significant cultural characteristics, strong economic benefits, and high mass participation, and it is relatively easy to form a brand, and the feasibility of industrialization is very high. Tibet Caisha Mandala Painting Project, Ewenki Clothing Project, New Fourth Army Shu Four Detachment Lightning Protection Project Site Project, Shaanxi Jiaxian Ancient Jujube Garden Project, and Gansu Gaolan Shilan Ancient Pear Garden Project belong to the category to be developed, with a score range of 50–70 minute. This type of project has the same advantages as the previous type of project but needs to be properly developed on the basis of better protection and inheritance. Emei Wushu Project and Wannian County Rice Farming Custom Project belong to the protection category, with a score range of <50 points. Due to the special nature of low economic benefits or due to the large degree of damage suffered by such projects, rescue and protection are more necessary than industrial development. It can be seen that the evaluation of the development value of the cultural heritage resources industry provides a basis for guiding the industrialization development and can effectively promote the development of cultural resources. In future research, the focus of follow-up research will focus on how to adapt to the regional conditions of industrial development, what kind of heritage protection mechanism to adopt, how to balance protection and development, and the impact of industrial development.

Data Availability

The data of this paper can be obtained upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this work.

References

[1] Y. Pang, W. Dong, and Y. Liu, "Research on the conservation and sustainable development strategies of modern historical heritage in the Dabie Mountains based on GIS," *Open Geosciences*, vol. 13, no. 1, pp. 717–735, 2021.

[2] B. Rouhani, "Cultural heritage in peacebuilding and economic development," *Antiquity*, vol. 92, no. 365, pp. 1397–1400, 2018.

[3] T. Yi, L. Zhou, C. Jing, J. Li, and X. Nie, "Integration of digital cultural heritage resources in China: understanding public expectations," *Libri - International Journal of Libraries and Information Services*, vol. 68, no. 1, pp. 59–70, 2018.

[4] C. Heitman, W. Martin, and S. Plog, "Innovation through large-scale integration of legacy records: assessing the "value added" in cultural heritage resources," *Journal on Computing and Cultural Heritage*, vol. 10, no. 3, pp. 1–10, 2017.

[5] F. J. Ruzic, "Identifying and ranking cultural heritage resources on geotagged social media for smart cultural tourism services," *Computing Reviews*, vol. 58, no. 8, pp. 494-495, 2017.

[6] X. Ding and Y. Tian, "Predicting retweeting behavior based on BPNN in emergency incidents," *Asia Pacific Journal of Operational Research*, vol. 34, no. 1, pp. 537–546, 2017.

[7] S. Li, T. Chen, L. Wang, and C. Ming, "Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index," *Tourism Management*, vol. 68, pp. 116–126, 2018.

[8] Y. Zhang, Y. Wang, H. Deng, X. Xiong, and H. Chen, "IAFSA-BPNN for wind power probabilistic forecasting," *Power System Protection and Control*, vol. 45, no. 7, pp. 58–63, 2017.

[9] P. Zhang, Y. Cai, and J. Wang, "A simulation-based real-time control system for reducing urban runoff pollution through a stormwater storage tank," *Journal of Cleaner Production*, vol. 183, pp. 641–652, 2018.

[10] G. E. J. Mogomotsi, P. K. Mogomotsi, R. Gondo, and T. J. Madigele, "Community participation in cultural heritage and environmental policy formulation in Botswana," *Chinese Journal of Population Resources and Environment*, vol. 16, no. 2, pp. 171–180, 2018.

[11] D. Diaz-Corona, J. Lacasta, M. A. Latre, F. J. Zarazaga-Soria, and J. Nogueras-Iso, "Profiling of knowledge organisation systems for the annotation of Linked Data cultural resources," *Information Systems*, vol. 84, pp. 17–28, 2019.

[12] P. C. Guzman, A. P. Roders, and B. Colenbrander, "Measuring links between cultural heritage management and sustainable urban development: an overview of global monitoring tools," *Cities*, vol. 60, no. 1, pp. 192–201, 2017.

[13] F. Rui, X. Romo, and E. Paupério, "Flood risk assessment of cultural heritage at large spatial scales: framework and application to mainland Portugal," *Journal of Cultural Heritage*, vol. 43, no. 1, pp. 163–174, 2020.

[14] P. Tarlow, "Can a multi-cultural industry such as tourism be inclusive?" *Worldwide Hospitality and Tourism Themes*, vol. 12, no. 6, pp. 657–662, 2020.

[15] Y. Zhao, "Digital protection of cultural heritage based on web technology," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3196063, 9 pages, 2022.

[16] J. Xie, "Innovative Design of Artificial Intelligence in Intangible Cultural Heritage," *Scientific Programming*, vol. 2022, Article ID 6913046, 8 pages, 2022.

[17] Z. Wei and Q. Tong, "Data-driven historical preservation: a case study in Shanghai," *Neural Computing & Applications*, vol. 32, no. 8, pp. 3423–3430, 2020.

[18] Y. Liu, "Application of digital technology in intangible cultural heritage protection," *Mobile Information Systems*, vol. 2022, Article ID 7471121, 8 pages, 2022.

Hindawi

Research Article

Trust-Based Smart Contract for Automated Agent to Agent Communication

Halima Mhamdi [ID],[1] **Ben Othman Soufiene** [ID],[2] **Ahmed Zouinkhi** [ID],[1] **Obaid Ali** [ID],[3] **and Hedi Sakli** [ID][1,4]

[1]*MACS Research Laboratory, National Engineering School of Gabés, Gabés University, Gabes 6029, Tunisia*
[2]*PRINCE Laboratory Research, ISITcom, Hammam Sousse, University of Sousse, Sousse, Tunisia*
[3]*Ibb University, Department of Computer Science & Information Technology, Tudun Wada, Yemen*
[4]*EITA Consulting 5 Rue Du Chant des Oiseaux, 78360 Montesson, Montesson, France*

Correspondence should be addressed to Obaid Ali; obaid.alii2016@gmail.com

Blockchain technology is now regarded as one of the most interesting and possibly innovative technologies. It enables information to be stored and exchanged securely and transparently without the need for a centralized authority to regulate it. Some of the primary benefits of this technology are the atomicity of the stored data. Given its features, this technology has the potential to provide answers to challenges encountered in a very sensitive sector, namely, Internet of Vehicles (IoV). In IoV, vehicles and service providers autonomously capture and produce data without human intervention. This exchanged data must meet certain criteria such as decentralization, automation, security, and stakeholder trust management. To overcome these challenges, the integration of blockchain technology and multi-agent systems is a key solution. Based on smart contracts, the proposed solution consists of exploiting role-based access control (RBAC) and attribute-based access control (ABAC) techniques. This solution removes the central authority (CA) to reduce maintenance costs and eliminate legacy threats from centralized systems. The results, obtained from consumption costs, show that the developed platform is characterized by security, availability, and privacy.

1. Introduction

With its evolution, the Internet of Things (IoT) is included in various sectors: healthcare, agriculture, energy, industry, and intelligent transportation systems [1]. It allows communication between different types of objects at any time and in real-time. In 2021, the number of connected objects was estimated at 35.82 billion and will exceed 75.44 billion by 2025 [2]. Included in Vehicular Ad-Hoc Networks (VANET), the vehicle becomes the connected object par excellence, forming a new paradigm called the Internet of Vehicles (IoV). In the IoV, vehicles become capable of establishing interactions with themselves and with other connected surrounding objects. This interaction is ensured thanks to the permanent Internet connectivity offered by the IoV. The vehicles will enable consumers to become familiar

with the Internet of Things through new services that bring more comfort, precision, and efficiency. Indeed, road safety and traffic fluidity as existing applications can be improved by this new concept. Other new services such as vehicle maintenance, parking access, fuel stations, and tolls are developed. As a result, vehicles and service providers autonomously capture and generate data without human intervention [3]. This exchanged data must meet certain criteria. The high connectivity of vehicles and connected objects in the IoV network facilitates malicious attacks. Therefore, the different parties involved in the communication process must interact in a secure and confidential manner. Let us also not forget that the centralized architecture of intelligent vehicle communication models can lead to an interruption of the IoV network if the central server fails. Therefore, decentralization of the system is a very

important challenge. To overcome these challenges of decentralization, automation, security, and stakeholder trust management in IoV, the fusion of blockchain technology and multi-agent system is a key solution.

Distributed ledger technology is often known as blockchain technology. In this peer-to-peer network, data is manipulated in an immutable, decentralized, and confidential manner. These characteristics increase the demand for the use of blockchain technology in various sectors. Included in the supply chain, the blockchain is involved in terms of data security and traceability [4]. Ethereum blockchain and the ERC20 interface are the basis of the proposed system. This system allows automating P2P exchanges, reducing the cost for a trusted third party by ensuring interoperability. Therefore, it guarantees the reliability and trust of information access. Combined with IoT, an extensive study exposed by Saxena et al. [5] and Bhushan et al. [6]. The developed IoT applications are characterized by privacy, confidentiality, and security of managed data. The authors in their article [7], with the same objective of ensuring security and privacy, have outlined the contribution of this technology in the design and development of the smart city. The integration of blockchain technology in IoV and health is the subject of several review articles [8–10].

Any real or abstract entity that can communicate with other entities in its environment is defined as an agent. The set of autonomous agents thus forms a multi-agent system (MAS). The collaboration between the components of the system allows for solving a predefined problem [11]. The MAS allows automating the interactions between the agents forming a fully decentralized system. Indeed, the MAS consolidate the efficiency and the confidence of the human/ machine or machine/machine communication. It also ensures the security of the agents. The authors [12] have exploited the multi-agent system and propose a trust management system between agents by using the technique of identification of agents via private and public keys. In the same context, ensuring security and trust in the multi-agent system shows relevant and reliable results in energy exploitation in a smart home, in logistics, and in an industry [13, 14].

Several researchers have combined the multi-agent system with blockchain technology to take advantage of its security and privacy features. Among the applications developed are carpooling systems, energy consumption, as well as food supply chain monitoring, and electronic voting [15].

Applied in the IoV, our proposed solution consists in designing a decentralized communication system between vehicles and vehicular service providers such as automatic fueling stations using both technologies: multi-agent system and blockchain. The main objective is to ensure security and trust between the agents of the system by establishing automated communication without human intervention via smart contracts. The focal contributions of this work are summarized as follows:

(1) We propose an autonomous agent-to-agent communication system.

(2) To reduce operating and maintenance cost and eliminate legacy threats from a centralized system, we combine the MAS technology with the blockchain technology.

(3) We propose a smart contract-based distributed and dynamic access control mechanism that enables system agents to share data securely.

(4) We implement our scheme on Ethereum blockchain. The performance evaluation, experimental results, and security analysis clearly prove the systems' efficient.

The rest of this paper is organized as follows: The second section is devoted to the study of existing works regarding the application of blockchain technology in IoV. The proposed system architecture is introduced in detail in Section 3. While Section 4 describes the simulation platform and the obtained results. The analysis and performance of the system is included in Section 5. Finally, Section 6 concludes this paper and gives some hints for further research.

2. Related Works

The IoV concept is based on vehicular communication. Vehicles are thus intelligent objects that share and exchange information with each other and/or with other connected objects. Vehicle-to-object (V2X) communication enables the creation of a multitude services for the intelligent transport ecosystem [16]. In V2X, security and trust are two essential elements because the manipulated data can be subject to attacks. To overcome this problem, many scientists have recently become interested in using blockchain technology in vehicular networks. Its integration with data trading makes it possible to facilitate P2P exchanges. Table 1 summarizes the several works presented in this section.

In this context, Chen et al. [17] proposed a framework based on a consortium blockchain that allows the exchange of data that is both secure and efficient between buyers and sellers. The developed solution is an iterative double auction method. It aims to maximize the welfare of the society while preserving the privacy of buyers and sellers so that more users are encouraged to take part in data trading. In addition, the cost of data transmission is also considered to improve the stability of the system. This system is composed of 3 entities, namely, vehicles, edge layer, and blockchain layer. Vehicles interact with each other to exchange data. To access specific types of data, any car in the system must pay a virtual token. Depending on its current state and data needs, the purpose of each node may change. Data brokers are edge servers in the edge layer of the IoV framework. Each car communicates its data needs to the nearest data broker, which then informs local providers. Blockchain, smart contracts, and miners are the three main mechanisms in the blockchain layer. The blockchain is used to ensure high credibility and security, smart contracts allow for a variety of user-designed algorithms, and mining ads to the resilience of the system. The edge layer and the blockchain layer can work together in two ways. For ledger storage and consensus, the edge layer provides significant storage and computational

TABLE 1: Summarized literature review.

References	Objectives	Blockchain			Performance		
		Type	Consensus algorithm	Smart contracts	Security	Privacy	Trust
[17]	Ensuring secure and efficient data exchange	Consortium blockchain	PoW	*	*		
[18]	To facilitate efficient IoV data-trading	Consortium blockchain	PoW	*	*	*	
[19]	Ensure trust and decentralized of incoming communications between vehicles and their neighbors	NM	PoW and PoS				*
[20]	Solve the driving endurance problem of an electric car by implementing a vehicle-to-vehicle energy exchange system	Fabric blockchain	PoA	*	*		
[21]	Ensure security and confidence in vehicle registration	Ethereum blockchain	dPoW	*			*
[22]	Provide security when transmitting messages between vehicles in the IoV network	NM	NM		*		
[23]	Detection and notification of vehicle accidents in an ITS environment	NM	PBFT		*		
[24]	Provide automation of data exchange processes in a secure way between the different participants of the IoV system	Ethereum blockchain	NM	*	*		
[25]	Overcome the scalability problem for trust management in the IoV system	Ethereum blockchain	PoW	*			*
This paper	Ensure security and trust between the agents of the system by establishing automated communication via smart contracts.	Ethereum blockchain	PoW and PoS	*	*	*	*

resources. To create trust and ensure security, the blockchain layer enables an edge layer.

Using blockchain technology, the authors in [18] offer a secure and decentralized IoV exchange system. This system uses an incentive-based investment and pricing structure, enabling car lending. By promising to pay interest and an incentive, a vehicle with a loan request can borrow from a group of cars. To optimize the payoffs for borrowers and lenders, a two-stage Stackelberg game is established. The perfect subgame equilibrium at each stage is analyzed by inverse induction. Furthermore, for secure data exchange and peer-to-peer financing services, the authors exploited the consortium blockchain, smart contracts as well as the proof of work as a consensus mechanism. The numerical results show that both the suggested independent and uniform pricing systems are effective.

Yang et al. [19] develop a blockchain-based decentralized trust management approach in vehicle networks. The Bayesian inference model is used by cars in this system to verify the signals received from nearby vehicles. They use the combination of proof-of-work (PoW) and proof-of-stake (PoS) consensus method. Through this fusion, all RSUs work together to keep the trust blockchain up to date, reliable, and consistent.

The study in [20] proposes a vehicle-to-vehicle (V2V) energy exchange mechanism to solve the driving resistance problem of electric vehicles. In the distributed BIoV with imperfect information sharing, Bayesian pricing is used. The proposed approach has several advantages over a static set with all available data. Electricity transactions are more

trustworthy, secure, and reliable using blockchain technology. For starters, the public/private key pair and digital signature are used with blockchain to address the security concerns of vehicle users. Second, the pricing method is incorporated into the smart contract for electric transactions, which not only reduces the dependency on the middleman, but also ensures that the transaction is fair.

Javaid et al. [21] address the problem of trust in the IoV network. They rely on blockchain technology to overcome this challenge by developing a trust protocol. The proposed solution uses both smart contracts and dynamic proof of work (dPoW) as consensus algorithm as well as certificates and nonclonable physical functions. These 4 mechanisms work together to ensure security and trust in vehicle registration. The resulting solution is characterized by throughput scalability, decentralization, low latency, and vehicle privacy protection. Similarly, to solve the problem of security during the transmission of messages between vehicles in the IoV network, a protocol named IoV-SMAP is implemented. This protocol makes it possible to overcome security flaws while guaranteeing the confidentiality and authentication mechanisms of the users.

In their paper [22], the authors propose a protocol for vehicle accident detection and notification in an ITS environment using a new authentication method called BCAS-VADN. They use blockchain technology to store the information generated by this method. In addition, BCAS-VADN offers end-to-end encryption and the ability to store a cryptographic hash signature of the interactions included in the blocks of the blockchain, thus providing a method of
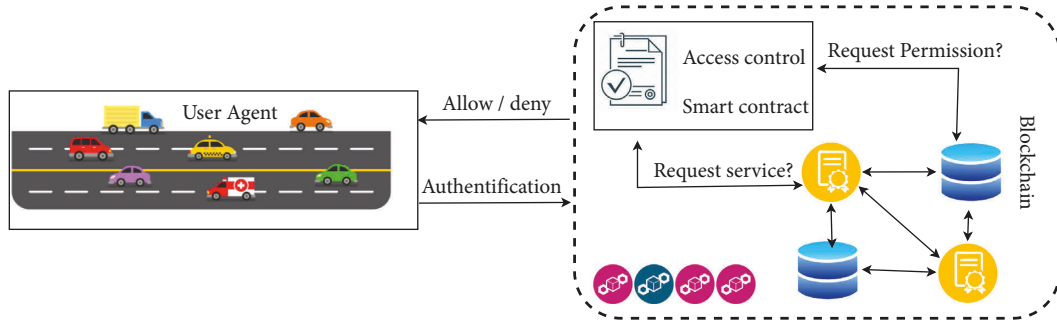
validating the integrity of transactions. The proposed system is composed of 4 phases: registration of all entities in the system, their authentication, verification, and addition to the blockchain and the dynamic addition of new nodes to the network. The use of blockchain technology and the Practical Byzantine FaultTolerance consensus mechanism reinforce the security since it is impossible to delete or modify the data exchanged in the network.

PSEV [23] is a framework based on blockchain technology that allows the creation of payments and secure communications between the different participants of the IoV system. This solution ensures the automation of data exchange processes using smart contracts. To validate the proposed solution, the authors use the Ethereum blockchain to develop a decentralized real-time parking reservation and payment application. The designed system is characterized by a low cost and a reduced execution time compared to existing systems. It also ensures integrity, immutability, and confidentiality.

Singh et al. [24] proposed a mechanism for trust management in the IoV system. The studied approach is to overcome the scalability problem by exploiting blockchain technology. The decentralized feature is ensured by using smart contracts and Proof of Work as a consensus mechanism. For validation, the authors resort to the use of the Ethereum blockchain platform that allows examining the performance of the system considering the average throughput and execution time. For the same purpose, Zhao et al. [25] also addressed the problem of trust management in the IoV system. They developed a protocol that bundles the practical Byzantine fault-tolerance consensus mechanism with the proof-of-work. This linkage ensures the anonymity of vehicles via identity-based group signature.

3. System Design and Architecture

The proposed system is a platform allowing the communication between autonomous agents by integrating the blockchain technology and the multi-agent system (MAS). The architecture of the system is composed of 2 main parts: a MAS that allows communication between agents and a public blockchain network (BC) that allows to store all transactions and smart contracts.

Blockchain technology, namely, smart contracts, allows agents to engage more quickly, automate business

procedures, and transmit orders across them. The data generated by multi-agent modeling can be kept in the blockchain. Agents are critical in IoV systems, and with blockchain, they may provide even more value to the ecosystem, increase its dependability, and facilitate its expansion.

3.1. System Design. Our system's model is depicted in Figure 1. Connected automobiles can communicate with a variety of items. They are also linked to the Internet through the provision of intelligent services. The most important feature is the decentralized exchange of information between different entities. The proposed model is composed of two actors or two agents: service providers and vehicles. All agents have their blockchain address. The blockchain records the transactions between the actors, as well as information about the vehicle and services. The smart contract ensures these transactions automatically validated by consensus mechanisms. The proposed model is based on 3 main concepts as follows:

3.1.1. Smart Contract. Smart contracts are the most important component of any blockchain framework as they fulfill basic functions. For the design of our framework, the first step is the deployment of different smart contracts either for system stakeholder's enrollment or for authentication to benefit from a service and pay the provider.

3.1.2. Authentication. Access to the system requires user authentication through Ethereum addresses for each agent. After authentication, the vehicle agent can consult and communicate with the service provider agent.

3.1.3. Access Control. Access control is a process that allows only authorized entities to manage information and controls this information. In our case, to use a service, the vehicle agent sends an access request via the smart contract that verifies the identity of the requester and authorizes him to send a request to the service provider.

3.2. Blockchain and Agent Communication Process. The communication process between our system's agents is divided into five stages: agent enrollment, agent registration,

Symbol	Description
A_X	Stakeholders agent in the communication process
VA	Vehicle agent
GSA	Gas station agent
S_L	Station localization
R_D	Remaining distance
M2M	Machine-to-machine

agent authentication, request service, and payment. The registration and communication process between agents based on the smart contract is explained by Algorithms 1–3 as described below. For convenient reference, Table 2 lists the various symbols and acronyms used in this subject.

3.2.1. Agent Enrollment. In this step, the system administrator assigns to each Ethereum address an agent type or a role. This designation is recorded in a hash table via a smart contract. The added role is then used when adding an agent or in the handling of transactions in the communication process. However, the deployment of different smart contracts required by our system was the object of this step.

3.2.2. Agent Registration. The registration step consists of adding the agents to the system. After verifying the account address, the agent can add its information to the Agent_DB through a smart contract. In this phase, the smart contract saves the characteristics or attributes of each agent, especially the identifier and its account in token. Upon successful registration, agents are allowed to join the blockchain.

3.2.3. Agent Authentication. To communicate and benefit from a service, the registered agents authenticate themselves. Two types of access control are used: Role-Based Access Control and Attribute-Based Access Control. The use of the "msg.sender" variable of the OpenZeppelin library allows for identifying and validating the agent's address. On the other hand, during the payment phase, it is necessary to control the access to the data of the agents such as the balance account.

3.2.4. Request Service. Once enrolled, the vehicle agent sends a request to the system to consult the list of available service providers. Then it chooses the appropriate provider. In this step, a smart contract is established between the stakeholders. This contract contains all the necessary criteria to use a service such as the minimum amount that must be in the applicant's portfolio.

3.2.5. Payment. After having exploited the desired service, a request is sent to the blockchain to complete the payment phase. The request containing the amount to be paid is managed automatically by a smart contract. This smart contract accesses the account of the stakeholder in this process and settles what is due.

3.3. Smart Contract Design. The suggested smart contract design comprises of three contracts: Registration_Contract, VA_SA_Contract, and Access_Contract.

3.3.1. Registration_Contract. It is designed to add Vehicle_Agent and Service_Agent in the system. The information provided by each agent is then stored in Vehicle_Agent_DB and Service_Agent_DB. These two registration tables represent a distributed database containing all registered agents. The authentication and authorization of agents are also controlled by the logging database.

3.3.2. VA_SA_Contract. It is a binding agreement between the vehicle agent and the service agent. It includes the system's stakeholder ID as well as the service charge.

3.3.3. Access_Contract. It maintains access control among system agents. When sending a service request from the vehicle agent to the service agent, this contract controls the access of this request. This contract is also involved in the payment phase. When access to the service is finished, using ABAC, it deposits the requested amount from the vehicle wallet to the service wallet.

4. System Implementation and Analysis

Figure 2 presents the system operation as well as the simulation tools used, which will be discussed later. The proposed platform consists of two agents: VA and GSA. The latter two form an autonomous system allowing the automation of "M2M communication" processes. First, VA, when it reaches a fuel threshold, will choose a GSA from a list of stations registered in the blockchain. This list provides information about the location of the station and its rates and its availability. The loading process requires the authentication of the vehicle requesting the service and the provider station. This is done via decentralized application interface (Dapp). BC is used to eliminate the middleman for authentication, payment, and automation of the service via smart contracts. After choosing the station, a smart contract is established between the two parties for the loading. The vehicle pays the loading amount which depends on the loaded quantity, type, and price, but the money is kept in the blockchain. When the loading process is completed, the smart contract automatically pays the agreed amount of money to the station account.

4.1. Simulation Platform. The agents of our system are modeled by the Netlogo Framework through which two types of communication are implemented between the different agents: the request and the reception of a transaction. The conversation between the agents is ensured by an agent communication language (ACL). To facilitate the communication between Netlogo, the Dapps, and the blockchain network, the Maven extension is used.

Each agent has an account in the Ethereum blockchain. To access this account, we use the web3.js library through an
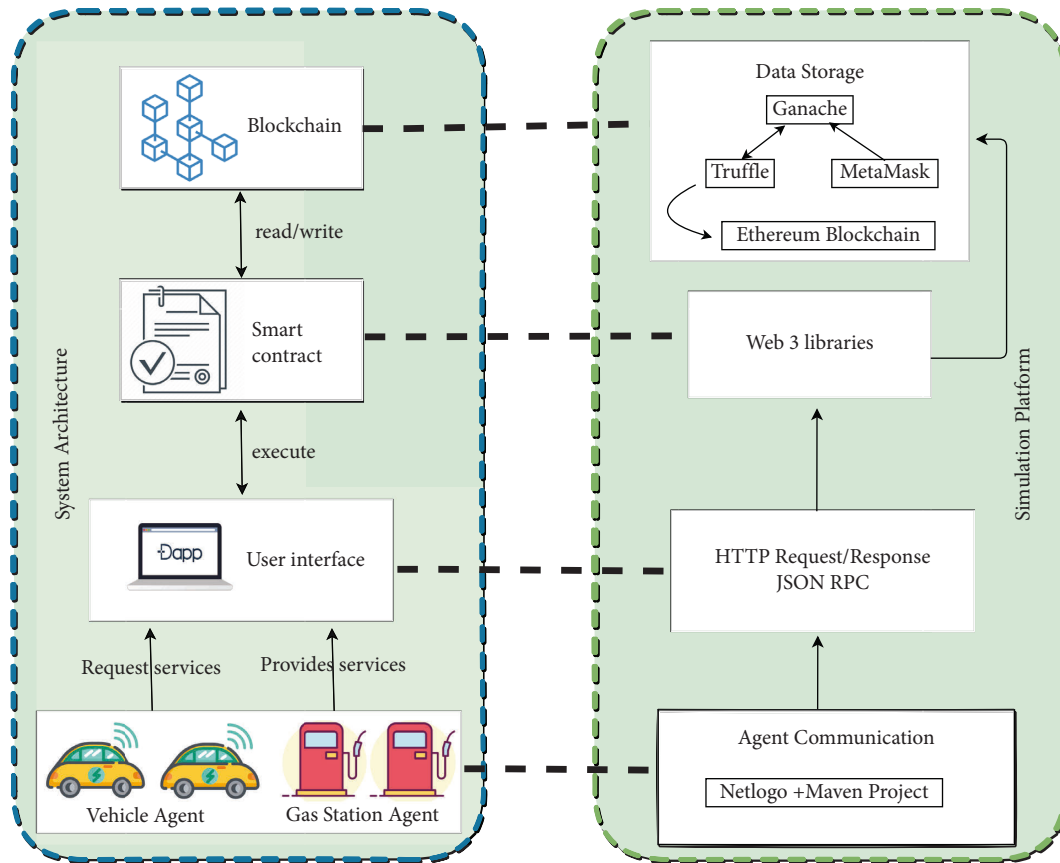
FIGURE 2: Proposed system diagram.
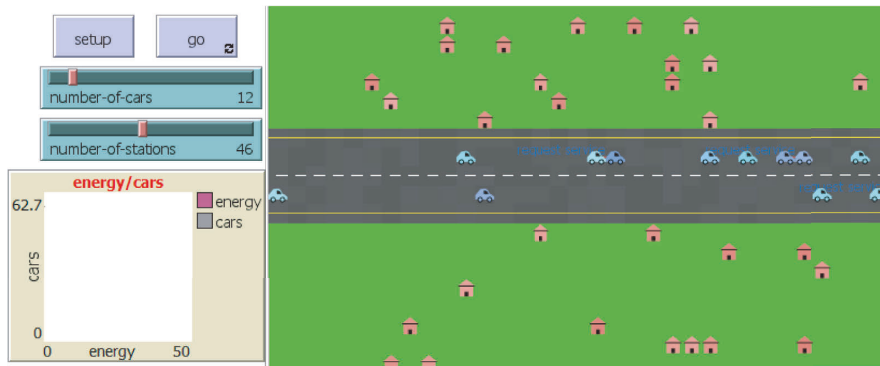
FIGURE 3: Netlogo interface.

Input: request to register from the smart contract
Output: agent added to the smart contract
(1) procedure assign_agent (Agent_type, Account)
(2) add agent and account in agent's mapping
(3) end procedure.

ALGORITHM 1: Agent enrollment.

HTTP connection in JSON RPC format. The truffle development environment is used to compile and migrate smart

contracts to the blockchain. Metamask is used to implement Ethereum wallet functions that allow participants to control

FIGURE 4: Smart contracts compilation.
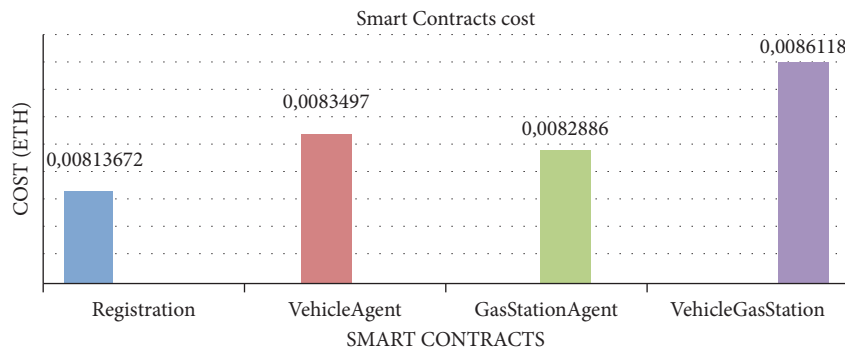
FIGURE 5: Smart contracts cost.

the Ethereum account information and make the transaction payment. Any changes submitted to the transactions are recorded on the blockchain network (see Figure 3).

Table 3 provides a list of the technologies used to implement the system and a brief description of each.

4.2. Results and Performance Evaluation. In this section, we will present the results of smart contracts deployment of our system as well as the consumption cost and transaction cost.

4.2.1. Smart Contracts Deployment. Figure 4 depict the compilation and migration of smart contracts to the blockchain ganache. Five smart contracts are compiled and deployed in the "development" blockchain: "Gas-StationAgent.sol," "Registration.sol," "Auth.sol," "VehicleAgent.sol," and "VehicleGasStation.sol." The entire cost of this transaction as shown by the migration result, is 0.003796388 ether, which is the equivalent of £ 13.21. This conversion is done on March 28th, 2022.

Figure 5 represents the execution costs of the different smart contracts in our system. We notice that the cost of the vehicle gas station contract is higher than the other contracts. This contract contains the main functions of the system, such as communication between the different agents as well as the access control functions ABAC and RBAC. While registration contract consumes less gas since it just assigns a role to the agents in a table containing the accounts and roles. Vehicle agent and gas station agent contracts take care of adding the relevant information to each agent. Their gas consumption costs are nearby.

Following the transfer of our smart contract, we will build a local virtual server containing the client-side application using the truffle framework. To join to our blockchain network, we need to connect to our Metamask portfolio as shown in Figure 6.

4.2.2. Consumption and Transaction Cost. The system's cost of consumption is calculated in terms of the number of gas units utilized by smart contracts. Gas is a unit of
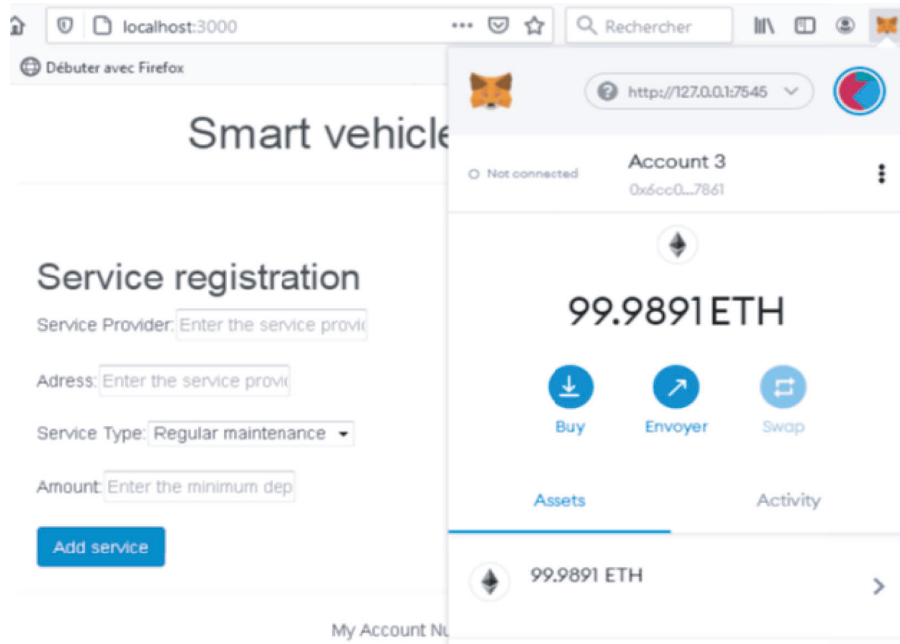
FIGURE 6: Vehicle registration data screen.

ALGORITHM 2: Upload agent information

measurement that estimates the processing power required for transaction execution. Gas units are transformed into ether value in Ethereum. To calculate the gas fees for each transaction in our system, we apply the following equation:

$$\text{Transaction Fee} = \text{gasUsed} \times \text{gasPrice}. \tag{1}$$

where gas used is defined depending on the storage and processing quantity for each transaction and gas price denotes the amount of Gwei necessary for the transaction. Let us take the example of the agent enrolment function which allows assigning to each agent a role and account. The gas used of this is 112883 and gas price is 20 Gwei. So,

$$\text{Transaction Fee} = 112883 \times 20 = 2257660$$
$$\text{Gwei} = 0.000225766 \, \text{Ether}. \tag{2}$$

We have calculated the transaction costs for the different smart contracts. Table 4 shows the gas consumed. Note that

the transaction cost for reading data from the blockchain, such as the "getVehicleAgent," "getAllGasStation" and "readMessage" functions, is null. Since mining is not required while getting messages from blocks, and no changes are required for the smart contract, this function does not incur any extra costs.

Figure 7 shows the gas consumed by each function in our system. The functions of the agent communication subsystem consume more gas with 0.000470134 Eth, 0.000297561 Eth and 0.000244863 Eth for the respective requestService, payment, and sendMessage functions. This high cost is due to the operations and the data they manipulate. On the other hand, the functions that manage the addition of agents: addVehicleAgent and addGasStationAgent consume about 0.0003.

4.3. Comparison. Table 5 compares the proposed method to other works in the field in terms of performance parameters
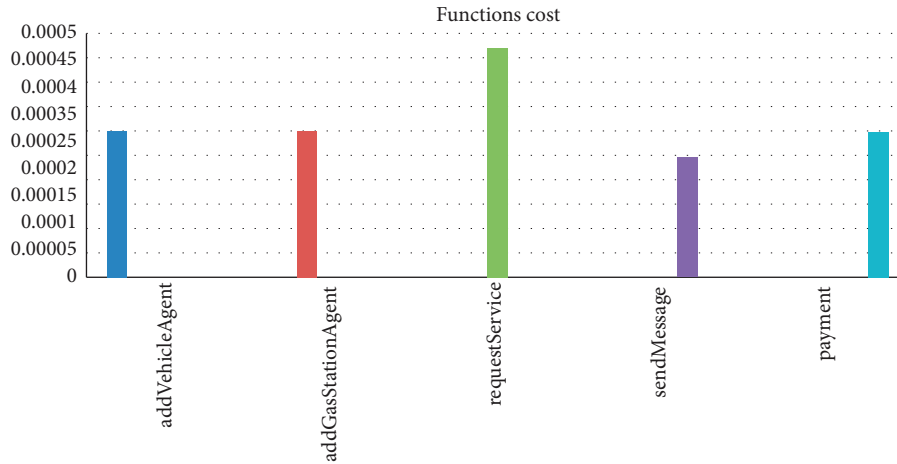
FIGURE 7: Functions cost.

```
Input: request to read data from the smart contract
Output: access to service provider
Initialization: connected to the application as a vehicle_agent
(1) function check service_data ()
(2) if (msg.sender == VA) then
(3) send request to blockchain
(4) if (S_L<=R_D) then
(5) return (station_liste)
(6) send bid request ()
(7) else
(8) return ("alert")
(9) end if
(10) else
(11) return unauthorized access
(12) end if
(13) end.
```

ALGORITHM 3: Search service provider.

TABLE 3: Technologies used for the realization of the system.

Technologies		Description
Blockchain	Ganache	A personal blockchain for local development. It allows to deploy smart contracts, develop applications, and run tests. Ganache provides 10 Ethereum accounts with a balance of 100 ether for each account, as well as a graphical interface to examine everything that happens in this blockchain.
	Truffle	A framework that provides a suite of tools for developing Ethereum smart contracts and a client-side interface, testing and deploying smart contracts in any Ethereum network.
	Ethereum	A distributed computing platform based on a public blockchain that can support advanced custom smart contracts using the turing-complete programming language.
	Smart contract: solidity	An object-oriented programming language dedicated to writing smart contracts. It is used to implement smart contracts on various blockchains, including Ethereum
Web server	Meta mask	A plugin allows to transform a web browser into a blockchain browser. It also allows the management of blockchain accounts, as well as ether funds to pay transactions
	Web3.js	A collection of libraries allowing interaction with a local or remote Ethereum node, using an HTTP or IPC connection in JSON RPC format.
Multi-agent system	Netlogo	A well-known and widely used cross-platform modeling and simulation environment for complex systems of simultaneously interacting agents, written on top of the Java virtual machine. Netlogo offers a wide range of generic features and operators to its users. In addition, to compensate for missing functionality, Netlogo is compatible with other platforms and libraries. Conversely, Netlogo can also be called and controlled by other programs. For this purpose, Netlogo provides a Java API.
	Maven extension	A program construction management system developed by the Apache foundation. It is based on the definition and use of POM (project object model) files containing all the instructions guiding the correct construction of the program.

Table 4: Function cost.

System	Function	Transaction cost (eth)	Price ($)
	Deploy contracts	0.003796388	13.21
	Agent enrolment	0.000225766	0.79
Vehicles agent management	addVehicleAgent	0.000332948	1.16
	getVehicleAgent	0	0
	getAllVehicleAgent	0	0
Gas stations management	addGasStationAgent	0.000301534	1.05
	getGasStationAgent	0	0
	getAllGasStation	0	0
Vehicles and gas station communication	Payment	0.000297561	1.03
	sendMessage	0.000244863	0.85
	readMessage	0	0
	requestService	0.000470134	1.63

Table 5: Comparison of proposed system with related work.

	[22]	[23]	[24]	[25]	Our system
Blockchain-based	Yes	Yes	Yes	Yes	Yes
Smart contracts	No	Yes	Yes	No	Yes
Access control-based	No	No	No	No	Yes
Multi-agent system	No	No	No	No	Yes
Security	Yes	Yes	No	Yes	Yes
Trustfulness	No	No	Yes	Yes	Yes

and used technologies. As previously stated, we were unable to locate a paper that shared our concerns. As a result, several works implemented a solution based on blockchain technology and smart contracts, but they did not use multi-agents' system. It is important to note that in the related works, we only focused on papers that are more relevant to our use-case and chose from them; as a result, some papers are not related to alternative access control methods.

4.3.1. Security. The use of the RBAC and ABAC mechanisms ensures the security of our proposed framework. So, no third party is allowed to access the system. Let us not forget also that blockchain is protected with mechanisms and protocols. Therefore, agent data can be handled reliably and confidentially. Only trustworthy persons have access to these data. The system denies access to any untrusted third party attempting to access the system.

4.3.2. Availability. The use of the public blockchain allows the system to be accessible to anyone who need a gas station service.

4.3.3. Trustfulness. Trust is maintained with access control via user registration as well as restricting access to the data of our system stakeholders.

5. Conclusions

In this study, we propose a smart contract-based distributed and dynamic access control mechanism that ensures an autonomous agent-to-agent communication system. As a result, blockchain technology is combined with multi-agent

systems and access control in order to guarantee security and trustfulness. To ensure that these features are implemented, various smart contracts are strategically placed. Vehicle Agent and Gas Station Agent are dedicated to the management of the agents' data. The smart contract vehicle Gas Station allows the exchange of data between vehicles and service providers and payment. The search for a neighboring service is handled by the multi-agent system. To evaluate the performance of the proposed system, the calculation of the costs of the different smart contracts and the costs of the functions that they contain is performed. The results obtained show that the developed platform is characterized by security, availability, and privacy.

In the proposed system, several data are manipulated and need to be stored securely so their storage is a challenge to consider when implementing a real system. So as a perspective cloud storage is considered. In this case, the data of the different agents as well as the smart contracts are stored in the cloud and the blockchain just contains their URL. Another perspective is the implementation of an access control mechanism for the access to the cloud.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] F. A. Almalki, S. Ben Othman, F. A. Almalki, and H. Sakli, "Eerp-Dpm: Energy efficient routing protocol using dual prediction model for healthcare using IoT," *Journal of Healthcare Engineering*, vol. 2021, Article ID 9988038, 15 pages, 2021.

[2] Y. Wu, H. Guo, C. Chakraborty, M. Khosravi, S. Berretti, and S. Wan, "Edge computing driven low-light image dynamic

enhancement for object detection," *IEEE Transactions on Network Science and Engineering*, vol. 1, 2022.

[3] S. B. Othman, F. A. Almalki, C. Chakraborty, and H. Sakli, "Privacy-preserving aware data aggregation for IoT-based healthcare with green computing technologies," *Computers & Electrical Engineering*, vol. 101, Article ID 108025, 2022.

[4] S. Hakak, T. R. Gadekallu, S. P. Ramu, P. K. R. Maddikunta, C. de Alwis, and M. Liyanage, "Autonomous Vehicles in 5G and beyond: A Survey," 2022, https://arxiv.org/abs/2207.10510.

[5] S. Saxena, B. Bhushan, and M. A. Ahad, "Blockchain based solutions to Secure Iot: background, integration trends and a way forward," *Journal of Network and Computer Applications*, vol. 181, Article ID 103050, 2021.

[6] B. Bhushan, C. Sahoo, P. Sinha, and A. Khamparia, "Unification of Blockchain and Internet of Things (BIoT): requirements, working model, challenges, and future directions," *Wireless Networks*, vol. 27, no. 1, pp. 55–90, 2020.

[7] V. Acharya, V. Ravi, T. D. Pham, and C. Chakraborty, "Peripheral Blood Smear Analysis Using Automated Computer-Aided Diagnosis System to Identify Acute Myeloid Leukemia," *IEEE Transactions on Engineering Management*, vol. 2021, pp. 1–14, 2021.

[8] A. K. M. B. Haque, B. Bhushan, and G. Dhiman, "Conceptualizing smart city applications: requirements, architecture, security issues, and emerging trends," *Expert Systems*, vol. 39, no. 5, 2021.

[9] H. Mhamdi, A. Zouinkhi, and H. Sakli, "Smart contracts for decentralized vehicle services," in *Proceedings of the 17th Int. Wireless Communications and Mobile Computing Conference (IWCMC 2021)*, Harbin, China, June 2021.

[10] H. Mhamdi, A. Zouinkhi, and H. Sakli, "Multi-agents' system of vehicle services based on Blockchain," in *Proceedings of the 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 291–296, Monastir, Tunisia, December 2020.

[11] H. Bhuyan, D. C. Chakraborty, S. Pani, and V. Ravi, "Feature and Subfeature Selection for Classification Using Correlation Coefficient and Fuzzy Model," *IEEE Transactions on Engineering Management*, vol. 2022, pub. ID. 3065699, pp. 2021–2115, 2022.

[12] N. Deepa, Q. V. Pham, D. C. Nguyen et al., "A survey on blockchain for big data: approaches, opportunities, and future directions," *Future Generation Computer Systems*, 2022.

[13] N. Sakli, H. Ghabri, B. O. Soufiene et al., "ResNet-50 for 12-lead electrocardiogram automated diagnosis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7617551, 16 pages, 2022.

[14] F. A. Almalki and B. O. Soufiene, "EPPDA: an efficient and privacy-preserving data aggregation scheme with authentication and authorization for IoT-based healthcare applications," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5594159, 18 pages, 2021.

[15] B. O. Soufiene, A. B. Abdullah, A. Trad, and H. Youssef, "Peerp: A priority-based energy-efficient routing protocol for reliable data transmission in healthcare using the IoT," in *Proceedings of the 15th International Conference on Future Networks and Communications (FNC)*, Leuven, Belgium, August 2020.

[16] A. R. Javed, M. A. Hassan, F. Shahzad et al., "Integration of blockchain technology and federated learning in vehicular (IoT) networks: a comprehensive survey," *Sensors*, vol. 22, no. 12, p. 4394, 2022.

[17] C. Chen, J. Wu, H. Lin, W. Chen, and Z. Zheng, "A secure and efficient blockchain-based data trading approach for Internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9110–9121, 2019.

[18] K. Liu, W. Chen, Z. Zheng, Z. Li, and W. Liang, "A novel debt-credit mechanism for blockchain-based data-trading in Internet of vehicles," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9098–9111, 2019.

[19] Z. Yang, K. Yang, L. Lei, K. Zheng, and V. C. M. Leung, "Blockchain based decentralized trust management in vehicular networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1495–1505, 2019.

[20] S. Xia, F. Lin, Z. Chen, C. Tang, Y. Ma, and X. Yu, "A bayesian game based vehicle-to-vehicle electricity trading scheme for blockchain-enabled Internet of vehicles," *IEEE Transactions On Vehicular Technology*, vol. 69, no. 7, pp. 6856–6868, 2020.

[21] U. Javaid, M. N. Aman, and B. Sikdar, "A scalable protocol for driving trust management in Internet of vehicles with blockchain," *IEEE Internet of Things Journal*, vol. 7, no. 12, Article ID 11829, 2020.

[22] A. Vangala, B. Bera, S. Saha, A. K. Das, N. Kumar, and Y. Park, "Blockchain-enabled certificate-based authentication for vehicle accident detection and notification in intelligent transportation systems," *IEEE Sensors Journal*, vol. 21, no. 14, Article ID 15824, 2021.

[23] R. Jabbar, N. Fetais, M. Kharbeche, M. Krichen, K. Barkaoui, and M. Shinoy, "Blockchain for the Internet of vehicles: how to use blockchain to secure vehicle-to-everything (V2X) communication and payment?" *IEEE Sensors Journal*, vol. 21, no. 14, Article ID 15807, 2021.

[24] P. K. Singh, R. Singh, S. K. Nandi, K. Z. Ghafoor, D. B. Rawat, and S. Nandi, "Blockchain-based adaptive trust management in Internet of vehicles using smart contract," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3616–3630, 2021.

[25] Y. Zhao, Y. Wang, P. Wang, and H. Yu, "PBTM: a privacy-preserving announcement protocol with blockchain-based trust management for IoV," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3422–3432, 2022.

Research Article

Control System Development and Implementation of a CNC Laser Engraver for Environmental Use with Remote Imaging

Hani Attar [ID],[1] Amer Tahseen Abu-Jassar [ID],[2] Ayman Amer,[1] Vyacheslav Lyashenko [ID],[3] Vladyslav Yevsieiev [ID],[4] and Mohammad R. Khosravi [ID][5]

[1]*Department of Energy Engineering, Zarqa University, Zarqa, Jordan*
[2]*Faculty of Computer Science and Information Technology, Ajloun National University, Ajloun, Jordan*
[3]*Department of Media Systems and Technology, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine*
[4]*Department of Computer-Integrated Technologies, Automation and Mechatronics,*
 Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[5]*Department of Computer Engineering, Persian Gulf University, Bandar Bushehr, Iran*

Correspondence should be addressed to Mohammad R. Khosravi; m.khosravi@mehr.pgu.ac.ir

This article is aimed at studying the features of the control systems development for a small-sized Computer Numerical Control (CNC) portative laser engraver. The CNC is implemented in mobile maintenance and repair platforms for remote sensing of the environment where the wild environment may not allow us to access the animals and places. The proposed work in this paper is based on recent research, which shows that applying the automated CNC speeds up the processes of repair, modernizes the equipment size, and significantly reduces the economic costs; accordingly, the authors developed a block diagram of a portable CNC laser engraver. The choice of the hardware was also made, taking into account the possibility of quick replacement in the field, which reduces the repair time and the cost of the developed layout. A control system based on the selected modules was synthesized, and a stability check was carried out using MatLab tools. To check the correctness of the developed control system, the authors developed and assembled an experimental layout to illustrate the results of engraving on such a layout. Finally, the stability and sensitivity of the proposed system have been obtained and proved that the system works in a comfortable zone of stability. The obtained results show that the proposed CNC laser engraver has achieved the expected improvements (high speed, small size, short production and repairing time, minimum human influence factor, and achieving a better outcome).

1. Introduction

The introduction of modern progressive technologies in all spheres of human activity was presented in the concept of the fourth industrial revolution (Industry 4.0) [1–5]; resulting in making it possible to apply and integrate modern information technologies such as Artificial Intelligence (AI) [6, 7], Big Data (BD) [8, 9], and Internet of Things (IoT) with hardware solutions based on Microcontroller Unit (MCU) [10–14]. Accordingly, it becomes possible to automate the production process through the introduction of new control systems such as Computer Numerical Control (CNC) machines or Programmable Logic Controller (PLC) [15–17]. The implementation of CNC and PLC provides a wide

opportunity to optimize the production process, improve the adaptation flexibility and customization for the release of new products, and minimize the influence of the human factor at the production stage. When combining the Computer-Aided Technologies (CAx) with CNC and PLC, the resulted symbiosis technique enables the implementation of the following solutions on the basis of CNC [18]:

(i) Additive 3D printing (plastic or metal)

(ii) Industrial milling (metals, plastic, or wood)

(iii) Laser engraving (metal, plastic, wood, etc.)

All of the above-mentioned solutions are regarded as an integral part of the Cyber-Physical Manufacturing Systems

(CPPS) that are used in enterprises with flexible production cycles.

During the analysis of the CNC machines' performance, the authors gave good attention to the NATO reference book and the Logistics Handbook, mainly for Logistics in the Maintenance and Repair sections. Indeed, understanding the repairing field is regarded as a set of measures to restore the suitability of the material part for the operation in the shortest possible time [19]. Armament Repair Shop Set (ARSS) is a one-sided sliding shelter of a repair shop that contains small-sized portable CNC machines, on the basis of which you can quickly repair or upgrade the equipment without sending it to manufacturers [20, 21]. Similar portable machines can also be useful for field work and small workshops. At the same time, it should be noted that most of the CNC machines that are mainly designed for industrial applications have large overall dimensions and cannot be used in ARSS systems [22].

Thus, the development of a small-sized portative CNC laser engraver with high engraving accuracy is an actual and promising research area. Accordingly, to conduct such a study, it is necessary to solve the following tasks:

(i) Conduct research and analysis of the vectors of input parameters for a laser engraver to obtain a good description

(ii) Establish the relationship between the input and output parameters for control system implementations

(iii) Develop a structural control system for laser engraver

(iv) Select hardware modules for the control system and executing devices, which rely on the selected hardware modules and synthesize the control system and the stability check that applies the theories of automatic control

(v) Assemble an experimental portable model of a laser engraver and conduct trial engravings based on the developed control system to confirm the main theoretical provisions of the proposed work and the performance of a portable laser engraver model

2. System Design and Implementation

Matching the selected CNC and the laser machine equipment's hardware is regarded as the backbone of the proposed work.

2.1. Match Selection of CNC Laser Machine Control Key Parameters and the Equipment's Hardware. The brightness of images taken with the help of a laser vestal from the CPC is directly deposited according to the cob parameters that are set at the input of the control system. Among the main parameters that directly affect the quality of engraving, as a rule, the following parameters should be distinguished [23]:

The number (density) of engraving lines, which depends on the quality of the original image;
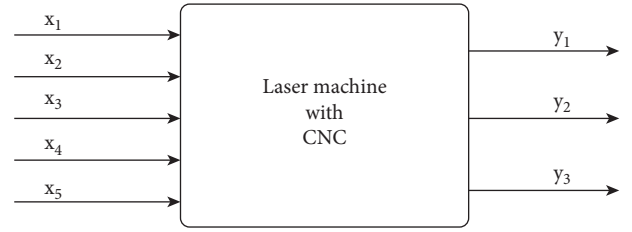
Figure 1: CNC laser machine as an object of control.

Laser power, which determines the choice of material for engraving and the maximum engraving depth;

The speed of the laser moving along the engraving surface and the engraving step determined by the type of Stepper Motors (SM) and the resolution of the laser beam, which depends on the type of the selected laser module, resulting in the need to configure the laser machine with CNC to ensure the quality of the image, represented as a "gray" box [24], with a vector of input parameters $x_1, x_2, x_3, x_4,$ and x_5, and the vector of output parameters $y_1, y_2,$ and y_3, as shown in Figure 1.

The vector of input parameters includes the number of image lines (lines/mm), power and speed of the laser (Watt), engraving step (mm), and laser beam resolution (dpi):

$$X = \{L, S, V, N, F\}, \tag{1}$$

where L is the number of lines (lines/mm), S is the laser power (Watt), V is the laser speed (m/s), N is the engraving step (mm), and F is the laser beam resolution (dpi). Depending on the requirements for the material on which the engraving is to be performed, the parameters F and S can be changed by installing a more powerful laser module. All other parameters can be changed programmatically through the CNC machine control system.

The vector of output parameters includes the line density per mm, the image engraving time, and the burn depth as shown in the equation as follows:

$$X = \{q, t, h, k\}, \tag{2}$$

where q is the line density (line/mm), t is the time (s), h is the burn depth (mm), and k is the image contrast. Note that the image contrast k depends on the movement speed of the laser module, which implements SMs, and the accuracy of the model position, which serves as the basis for choosing the type of SM.

The analysis of the relationship between input and output parameters of the control object allows drawing the following conclusions in the presented work:

(i) Number of lines L allows to achieve high-quality images by passing the laser a specified number of times

(ii) The laser power S sets its range of work associated with the speed of movement to simulate the color of the image, i.e., the depth of burning h

(iii) The laser movement speed V depends on the hardware of the engraver, which affects the time of engraving the image t
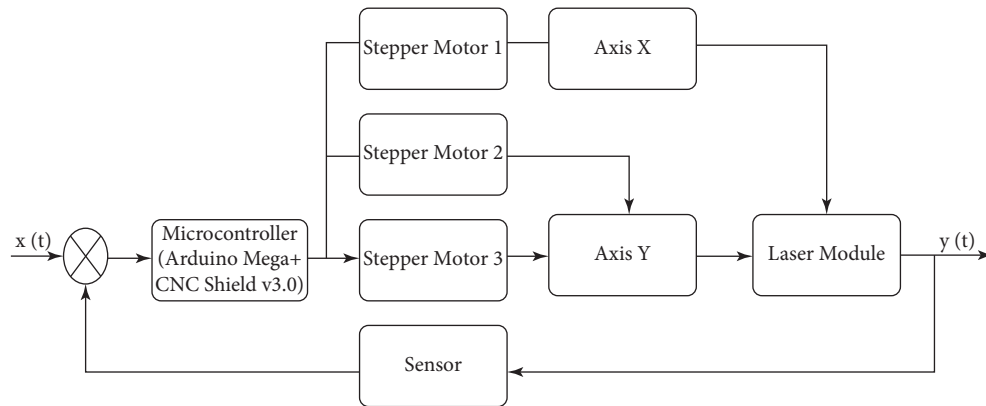
FIGURE 2: The structure of the control system of the CNC machine.
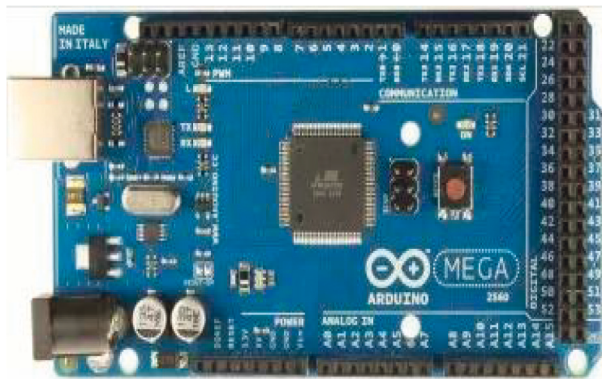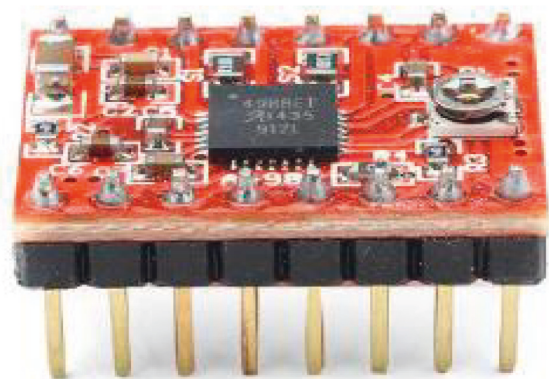
(a)   (b)

FIGURE 3: (a) Control board is Arduino Mega 2560-M16U [26]. (b) A4988 SM driver [27].

(iv) The image contrast k depends on the speed of the laser V

(v) The engraving step N depends on the density of the obtained lines q, that is, the fill density of the image

(vi) The detailed level of the received image, i.e., the quality of execution of an engraving, depends on the resolution of a laser beam

To create more control through CNC machines, it is necessary to select components such as control boards, SMs, temperature sensor, laser, and actuator (lead screws).

Based on the analysis of the relationship between the input and output parameters of the portable laser engraver developed, the authors propose the following structure of the cutting system with a laser engraver based on a CPC, as shown in Figure 2.

The input parameters are denoted by $x(t)$, which is a file in Bitmap Picture format (∗.bmp). The file ∗.bmp formatted file can be loaded into the control system, such as in microcontroller ATmega2560 [25], installed CNC Shield v.3.0 [26], and A4988 drivers [27], where 3-SMs are also connected to the board. SM 1 is responsible for moving along the X axis (Axis X), and SM 2 and 3 are responsible for moving along the Y axis (Axis Y). The use of two SMs allows ensuring the accuracy of moving the laser module along the

machine frame without using a more expensive SM with two shafts, resulting in reducing the cost of the experimental layout. The proposed solution allows moving the laser in Cartesian coordinates and controlling the engraving working area using the feedback through the Sensor as shown in Figure 2.

Figure 2 illustrates the way to select the elements of the machine control system with CNC, namely the microcontroller, SMs, actuators, and laser.

The selected control board in the proposed work is Arduino Mega 2560-M16U as shown in Figure 3(a), which has a low cost, simple instruction, and easy to provide a good interface and compatibility with operating systems, such as Windows, Macintosh OSX, and Linux, taking into consideration that most of the operating systems are focused only on Windows [28].

The selection of the Arduino Mega 2560 control board requires choosing an expansion board that reduces the number of wires and facilitates the installation. Drivers for the SMs are installed on the expansion board as in Figure 3(b), where jumpers for dividing the motor pitch are installed, which further affects the image clarity parameter.

CNC Shield v3.0 [26] was chosen as the expansion board, which is a specialized board for working with CNC machines as shown in Figure 4(a).

(a)   (b)
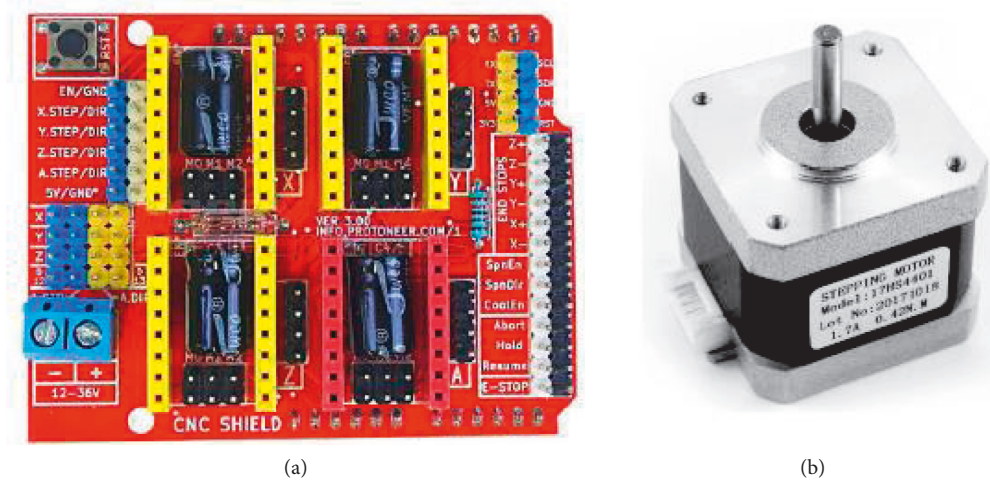
Figure 4: (a) CNC shield extension board v3.0 [26]. (b) Stepper mottor nema 17 [29].

(a)   (b)
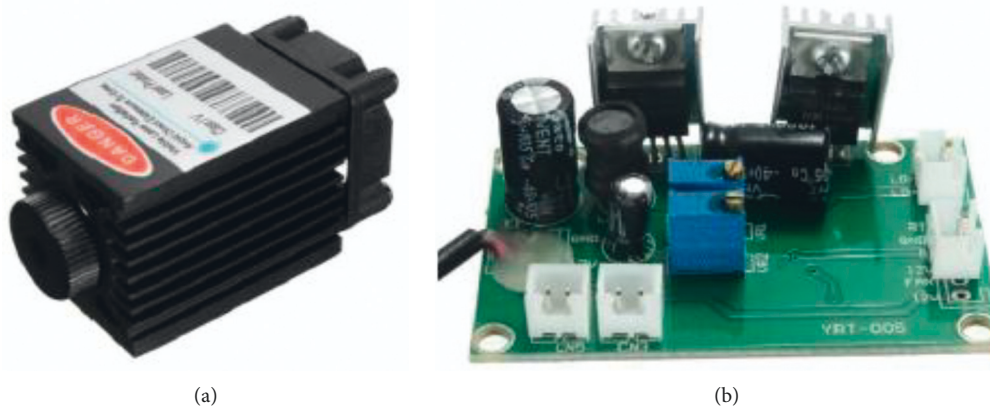
Figure 5: Focusable 500 mw diode laser and driver board [31]. (a) Low-power diode laser; (b) laser driver board.

Given that this small-bore portable laser engraver is being developed to be used in ARSS, the main condition is high maintainability and availability. Based on this, the authors believe that in the field, if one of the elements of the control system fails, it is possible to make an emergency replacement in the PnP (Plug and Play) mode. Therefore, the proposed hardware modules for the implementation of the control system are the best solution, in contrast to all the hardware elements of the control system integrated into one board.

Next, the SM is chosen, which converts electrical voltage pulses of the control voltage into discrete angular or linear movements of the rotor with its potential fixation in the desired positions. SM allows moving the functional tool throughout the working area of the machine with CNC. The functional tool is regarded as one of the main parts of the machine because it is responsible for the main operation, and where the CNC is intended for the surface treatment of the workpiece, or other operation. The hybrid Stepper Mottor Nema 17 (17HS4401), which is shown in Figure 4(b), is used for the control task of X and Y axis movement [29].

Powerful lasers are used on industrial CNC machines, the most common of which are CO_2 lasers and fiber lasers [30]. Such lasers require expensive maintenance. To decrease the maintenance cost, a low-power diode laser shown in Figure 5(a) and a laser driver board shown in Figure 5(b) [31] were chosen in the proposed work because it is easy to operate and has relatively high performance characteristics. Moreover, on the laser driver board, there is a TTL channel that monitors the power, where the laser is tuned off (zero power) at 5 V and tuned on (full power) at 0 V. TTL enables controlling a few parameters as the shades of gray, power of the laser module 1.6 W, wavelength 445–450 nm, color blue, supply voltage DC 12 V, and Frequency TTL from 0–20 kHz.

Comparisons of the studied laser parameters for small machines are presented in Table 1.

As can be seen from Table 1, the selected laser at an output power of 500 mW has different wavelengths and prices. Considering that a laser is a module with a life cycle of 3,000 to 8,000 hours of operation, the authors of the experimental machine with CNC under development chose the Focusable 500 mW, laser model.

TABLE 1: Comparison of the parameters of the studied lasers.

Laser model	Wave-length (nM)	Output power (mW)	Laser shape	Reverse polarity protection	Warm-up time	Price ($)
Focusable 500 mw [31]	808	500	Dot (rectangular or oval)	+	0	26
405MD-500 [32]	405	500	Dot (oval)	+	0	42
SRBI FB04-500 [33]	405	500	Dot (rectangular or oval)	+	0	41
FB03-500 [34]	405	500	Dot (oval)	+	0	62

2.2. *The CNC Machine Control System and the Generalized Main Parameters.* According to the selected control system elements of the CNC machine and the developed block diagram proposed in Figure 2, it is further necessary to synthesize the control system for each element. To do so, it is proposed to use the theory of automated control, where all the selected elements in the form of transfer functions are presented in this paper. The transfer function is one of the ways to mathematically describe a dynamic system. In automatic control theory, the transfer function of a continuous system is the ratio of the Laplace transform of the output signal to the Laplace transform of the input signal at zero initial conditions [35]. Accordingly, the developed control block diagram of the workbench with the CPC is described mathematically for each element through the transfer function [36].

The transmitter transfer function equation is

$$W_{\text{sen}} = \frac{0.4412}{0.02s^2 + s}. \tag{3}$$

The transfer function of the first engine equation is

$$W_{d1} = \frac{1}{0.1s + 1}. \tag{4}$$

The transfer function of the second engine equation is

$$W_{d2} = \frac{1}{4.49s + 1}. \tag{5}$$

The transfer function of the third engine equation is

$$W_{d3} = \frac{31.62}{0.4732s + 1}. \tag{6}$$

The transfer function of the laser module equation is

$$W_l = \frac{4.49s + 1}{0.000398s^2 + 0.01995s + 1}. \tag{7}$$

The transfer function of the first shaft equation is

$$W_{sh1} = 0.0335s + 1. \tag{8}$$

The transfer function of the second shaft equation is

$$W_{sh2} = 0.01s + 1. \tag{9}$$

Based on the developed structure presented in Figure 2 and the transfer functions shown in equations (3)–(9), the control system can be represented as the structure of the transfer functions, as illustrated in Figure 6.

The next step is to calculate the transfer function of the open and closed loop system [36, 37]:

The transfer function of the open loop system is given in the following equation:

$$W_p(s) = W_{mk}(s) \times ((W_{d2}(s) + W_{d3}(s)) \times W_{sh2}(s)$$
$$+ W_{d1}(s) \times W_{sh1}(s)) \times W_l(s) \times W_{\text{sen}}(s). \tag{10}$$

The transfer function of the closed-loop system is given in the following equation:

$$W_3(s) = \frac{W_{mk}(s) \times ((W_{d2}(s) + W_{d3}(s)) \times W_{sh2}(s) + W_{d1}(s) \times W_{sh1}(s)) \times W_l(s)}{1 + W_{mk}(s) \times ((W_{d2}(s) + W_{d3}(s)) \times W_{sh2}(s) + W_{d1}(s) \times W_{sh1}(s)) \times W_l(s) \times W_{\text{sen}}(s)}. \tag{11}$$

Making the substitution equations (1)–(7) into equations (8) and (9), the overall transfer function of the open loop system becomes the following equation:

$$W_p = \frac{2.098s^3 + 27.72s^2 + 68.71s + 13.95}{1.691 \times 10^{-6}s^7 + 0.0001902s^6 + 0.01062s^5 + +0.3212s^4 + 2.821s^3 + 5.103s^2 + s}, \tag{12}$$
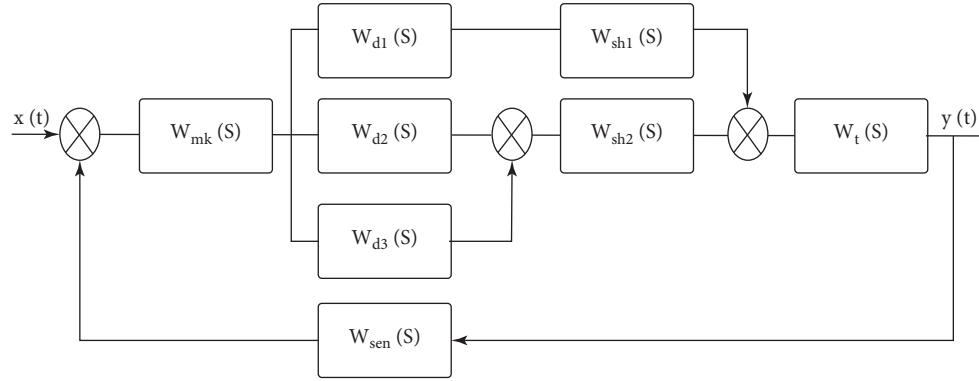
FIGURE 6: Structure of the automated control system of the CNC machine.

and of the closed loop system equation as follows:

$$W_3(s) = \frac{2.098s^3 + 27.72s^2 + 68.71s + 13.95}{1.691 \times 10^{-6}s^7 + 0.0001902s^6 + 0.01062s^5 + 0.3212s^4 + +4.922s^3 + 32.82s^2 + 69.71s + 13.95}. \tag{13}$$

The next step is to assess the stability of the open loop system at the roots of the characteristic equation. To do so, it is necessary for the transfer function of a closed loop system (13) to select the denominator in the form of a characteristic equation (14) and find its roots equation (15). If all roots of the characteristic equation have the same sign (either all negative or all positive), then the system will be stable. Otherwise, the system will not be stable.

$$1.691 \times 10^{-6}\lambda^7 + 0.0001902\lambda^6 + 0.01062\lambda^5 + 0.3212\lambda^4$$
$$+ 4.922\lambda^3 + 32.82\lambda^2 + 69.71\lambda + 13.95 = 0, \tag{14}$$

$$\begin{cases} \lambda_1 = -25.3646 + 34.9229\,i; \\ \lambda_2 = -25.3646 - 34.9229\,i; \\ \lambda_3 = -24.0819 + 3.6019\,i; \\ \lambda_4 = -24.0819 - 3.6019\,i; \\ \lambda_5 = -10; \\ \lambda_6 = -3.3537; \\ \lambda_7 = -0.2227. \end{cases} \tag{15}$$

Equation (15) shows that the sign of all roots are the same (negative), which means that the system is stable. To plot the location on the complex plane of the roots of the characteristic level equation (15), a package of applications is used to solve the technical calculation challenges by MatLab. Based on the stability conditions for linear automatic control systems, which can be formulated as follows: the system is regarded as stable when all the roots of its characteristic equation are located on the left side of the axis, i.e., all the real parts of the roots are negative, taking into consideration that it is enough to have one real part of any root in the right side of the axis to make the system unstable. As can be seen in Figure 7, the location of characteristic polynomial roots

on the complex plane is along the axis y, where y is the imaginary part, and along the axis x, where x is the real part. Accordingly, all roots are located on the left side of the axis. Therefore, the system is regarded as stable.

As a result, it is possible to assume that the closed loop system is stable; however, it is required to estimate the stability of the closed-loop system by Mikhailov's criterion to fully evaluate the proposed system, where the Mikhailov criterion is described as follows.

Mikhailov curve vector $D(jw)$ tests the system's stability by increasing the frequency from 0 to ∞, ∞, where the system is considered stable when it starts moving on a real positive semiaxis and rotates only counterclockwise, passing sequentially n quadrants of the coordinate plane, where n is the degree of the characteristic polynomial, i.e., the number of the roots of the characteristic polynomial [38].

The Mikhailov curve for stable systems is characterized by having a smooth spiral shape, and its end goes to infinity in that quadrant of the coordinate plane. If, as a result of the calculation, the control system of a portable laser engraver is stable, then the system becomes stable from external disturbances and will return to a state of rest.

Using MatLab, the Mikhailov curve was constructed, which is shown in Figure 8.

Though Figure 8 shows that the Mikhailov curve consistently covers 7 quadrants, to obtain clearer curves, the range of the coordinate axes can be changed as shown in Figures 9(a) and 10(b).

The graphs of Figures 9(a) and 10(b) show that the Mikhailov curve consistently covers 7 quadrants, which is equal to the degree of the characteristic polynomial without intersecting the origin, and fully meets the requirements of the Mikhailov criterion. Consequently, the closed loop system is regarded as stable.
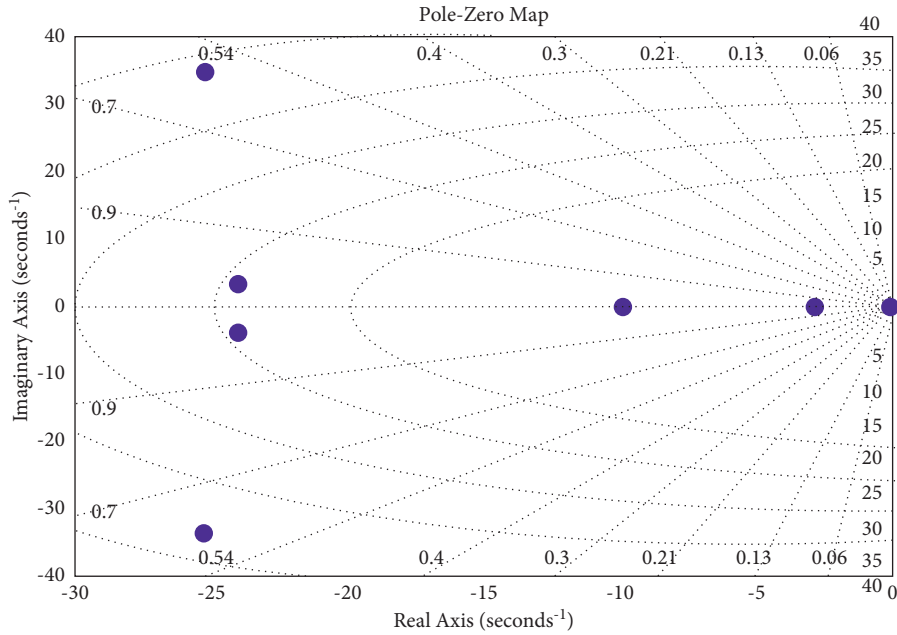
FIGURE 7: Location of the characteristic equation roots.
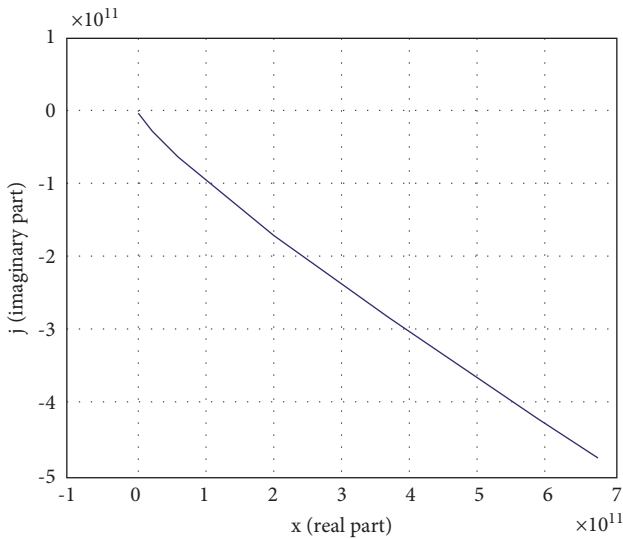
FIGURE 8: Mikhailov curve for the proposed system with 7 roots of the characteristic polynomial.

To define direct indicators of system's quality, the graph of the transient characteristic of the system is obtained and shown in Figure 11.

From Figure 11, the time of the system's regulation is determined, which is equal to $t_p = 0.805s$, and the value of over-regulation is calculated in the equation as follows:

$$\sigma = \frac{h_{max} - h_\infty}{h_\infty} \times 100\% = \frac{1.06 - 1.01}{1.01} \times 100\% = 4.95\%. \quad (16)$$

Under the operating conditions, the parameters of the control system vestal with CNC may vary within certain limits (wear, temperature fluctuations, friction, surface angle, etc.), where these fluctuations in parameters can lead to a loss of stability of the system mainly when operates close to the stability limit. Therefore, the designed ACS for CNC should work far from the limit of sustainability. In addition, it is necessary to determine the stability margins of the Westat control system with CNC by the values of the phase and amplitude margins; accordingly, a decibel log frequency characteristic (DLFC) and Logarithmic amplitude-phase frequency response (AFPFR) are obtained and shown in Figure 10.

Figure 10 identifies the following indicators of stability:

(i) The amplitude stability margin (ΔA) = 8.73 dB

(ii) The phase margin of stability ($\Delta\phi$) = 130°

The obtained values of stability indicate that the value of the margin of stability in the amplitude falls in the range of 6–20 dB, whereas the values of the margin of the phase fall in the range of 30°–600°, which indicates that the developed system is normally damped system. Based on the study of the stability of the developed control system of the CNC machine, the quality indicators and the stability margins, it is concluded that the selected elements of the system and its structure meet all the conditions of the design of such systems.

2.3. Optimization of CNC Machine Tool Control System.
CNC laser machine is a complex multidimensional object, and the values of its adjustable and outrageous parameters are dynamic. There are some difficulties in developing a management system for such facilities. For example, in the transfer function of the laser power control area that is applied to obtain the desired color intensity, which is calculated according to known methods of setting standard regulators for this class of objects, accordingly, corrections and, in many cases, nonstationary noise and perturbations acting on objects are required, resulting in considering the automatically

(a)

(b)
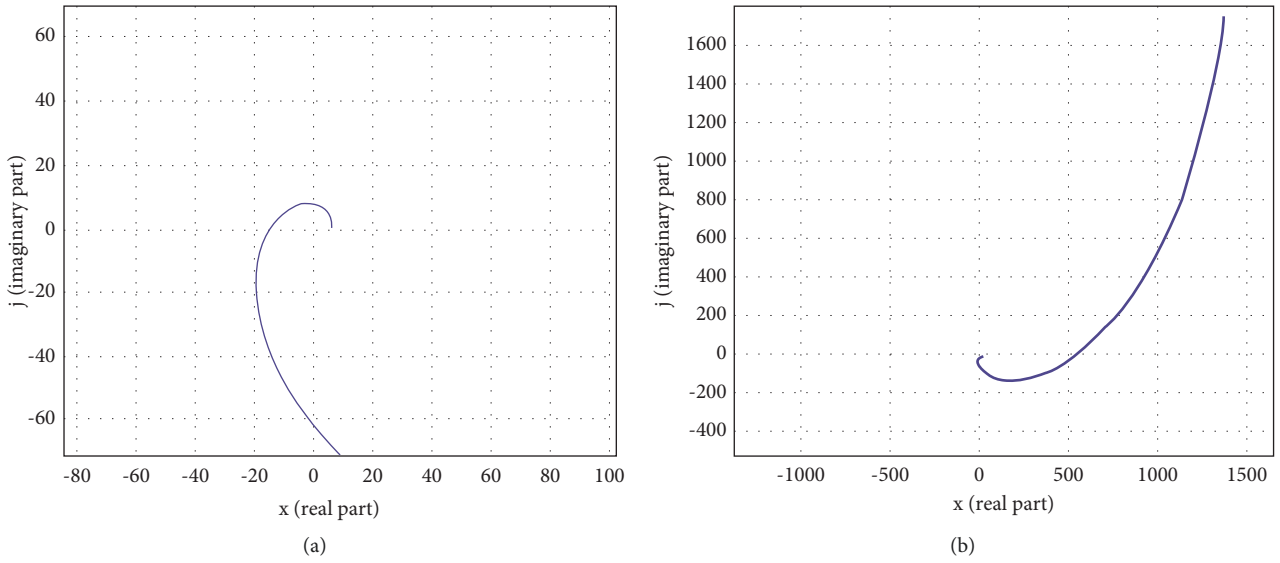
Figure 9: Mikhailov curve at the approximation range. (a) Range one. (b) Range two.
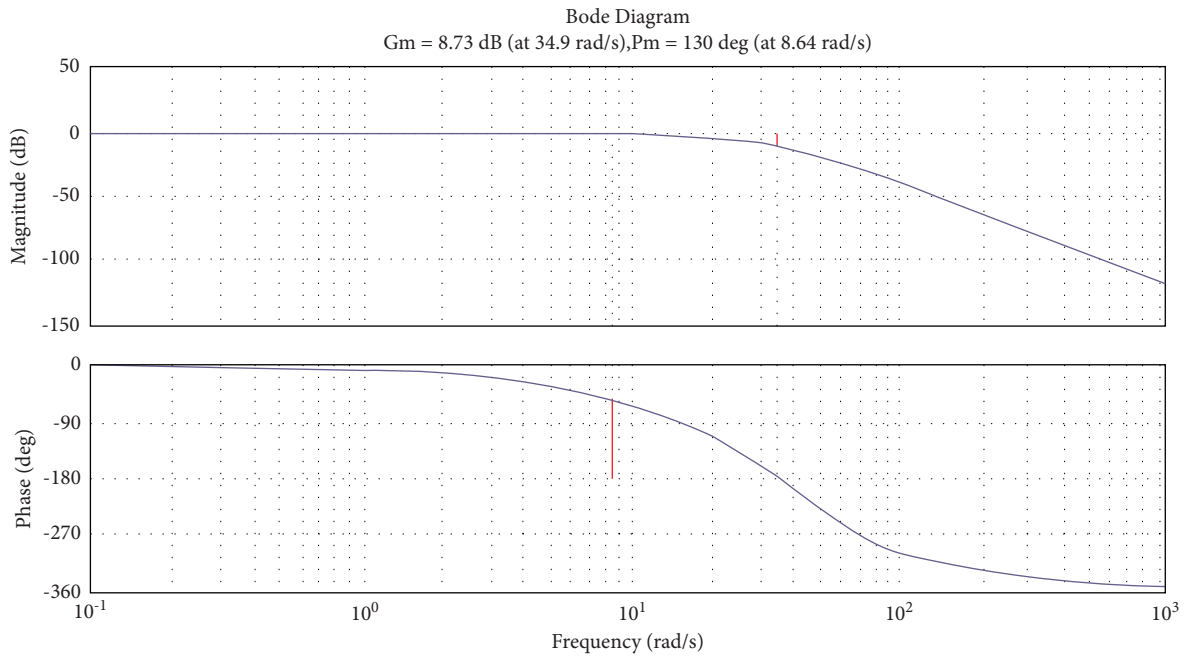
Figure 10: DLFC and AFPFR of CNC control systems.

adjustment of the control settings or selecting the optimal settings as an urgent task. For such a task, modeling is carried out by firstly building a model of the CNC machine control system with PID controller, as shown in Figure 12.

Determining the optimal parameters of the controller is an iterative process, which consists of 21 iterations, in which the parameters of the PID controller were selected and tested experimentally. At the same time, it is noted that the number of iterations depends on the requirements of the laser engraver control system (quality, reliability, accuracy, est.). The results of the optimization are the calculated optimal parameters of the PID controller that is shown in Figure 12.

Moreover, the graphs of the calculated transient characteristics are shown in Figure 13.

The obtained results shown in Figures 13 and 14 are not optimal because the parameters of the laser machine system controller are optimized according to certain parameters. To obtain optimal results for the system according to the operation parameters, a block diagram of the model based on Figure 12 is built, and the "Saturation" block is added, which is a nonlinear device that acts as an ideal limiter.

The saturation block keeps changing the output signal till it becomes equal to the input signal, which is the adjusted threshold for the upper and lower bounds. When reaching
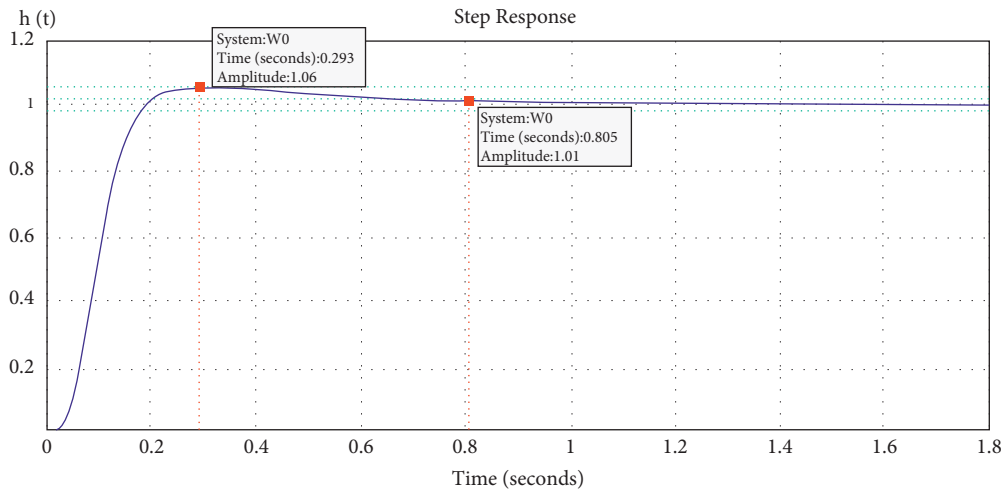
FIGURE 11: Transient characteristics of the CNC control system.

$$W_C(s) = \frac{2{,}098\,s^3 + 27{,}72\,s^2 + 68{,}71\,s + 13{,}95}{1{,}691{\cdot}10^{-6}s^7 + 1{,}902{\cdot}10^{-4}s^6 + 0{,}01062\,s_5 + 0{,}3212\,s^4 + 4{,}922\,s^3 + 32{,}82\,s^2 + {+}69{,}71\,s + 13{,}95}$$
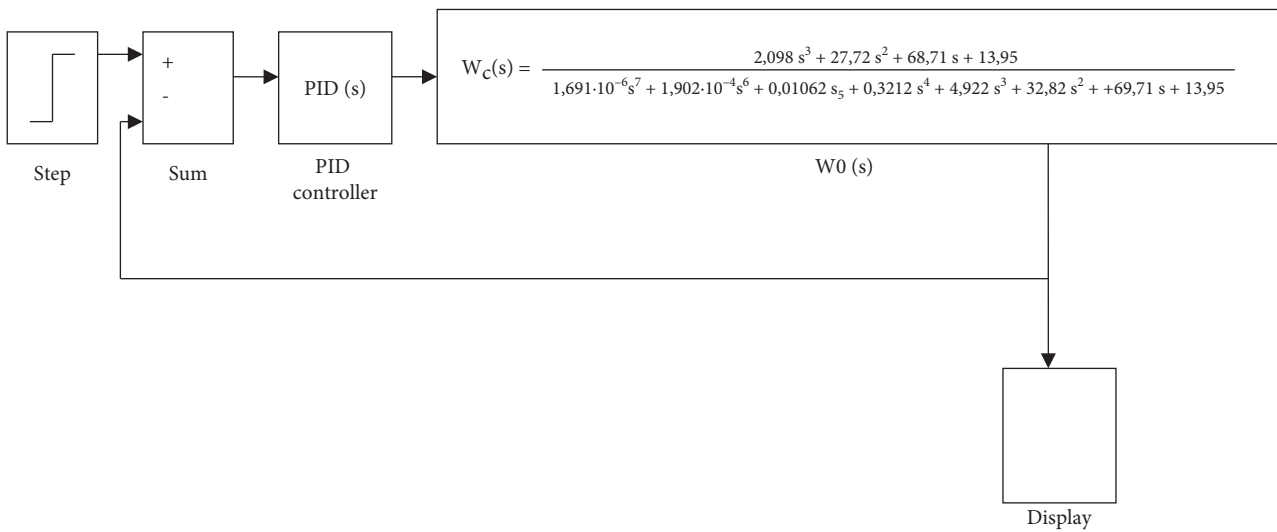
FIGURE 12: Layout of the CNC machine tool control system with PID controller in MatLab.
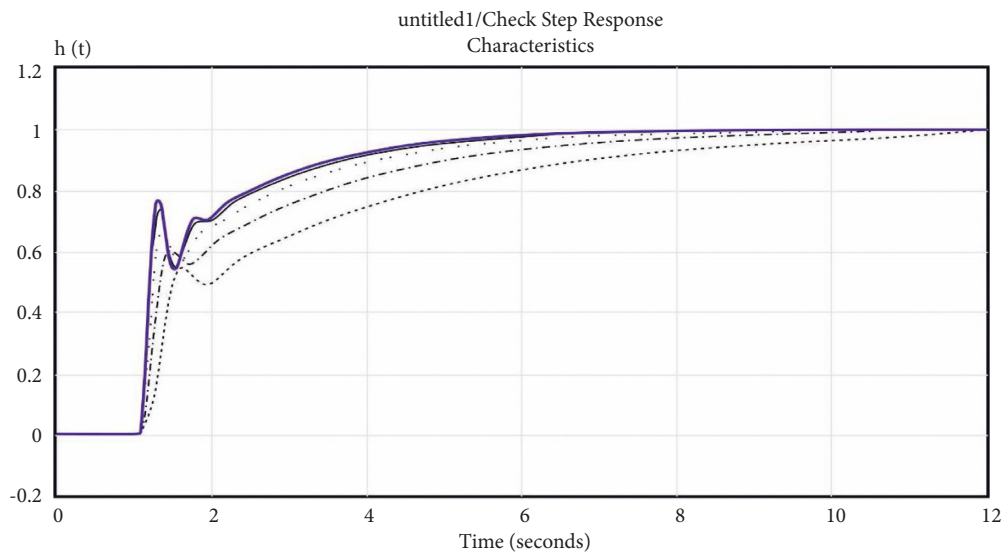
FIGURE 13: Simulation results of the transient characteristics.

DesignVars1(1,1)=   DesignVars1(2,1)=   DesignVars1(3,1)=

Name: 'Kd'   Name: 'Ki'   Name: 'Kp'

Value: -2.3303   Value: 87.8730   Value: 79.0469

Figure 14: The obtained values of the coefficients of the PID controller.
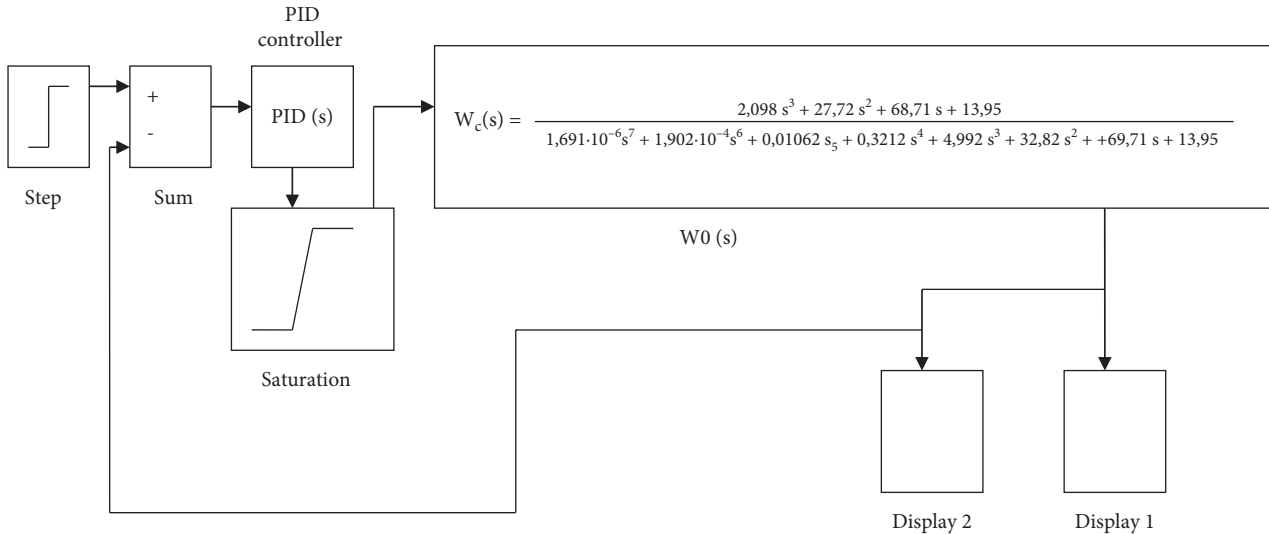
Figure 15: The investigated structural scheme of the model for optimization of the PID controller in MatLab.

Val (1,1)=   Val (2,1)=   Val (3,1)=

Name: 'Kd'   Name: 'Ki'   Name: 'Kp'

Value:79.6   Value: 87.32   Value: -4

Figure 16: Optimal coefficients of the proposed PID controller.

the adjusted thresholds, the output signal stops changing, resulting in improving the PID controller. Figure 15 shows the block diagram of the optimized model of the PID controller in MatLab.

Based on the simulation results, the PID controller parameters were optimized, and the optimal PID controller coefficients were obtained, which are presented in Figure 16.

The following coefficients were selected for the developed CNC workbench control system: coefficient of integral component = 87.32; coefficient of differential component = 79.6, and the coefficient of the proportional component = −4.

The transition graphs with optimized controller parameters are shown in Figure 17.

3. Development of the Structure and Prototype of the Laser Workbench Model with CNC

The full model structure contains hardware structure, parameters setting, and the operating software.

3.1. Hardware and Setting Development. Based on the proposed block diagram of the machine with CNC (Figure 2)

and the selected hardware modules (Figures 3–5), a block diagram was developed, which is presented in Figure 18.

Accuracy is important for CNC, so the design with the use of profile and rails allows for achieving the highest possible accuracy, which determines the scope of the application. The surface of the rail is pretreated and carefully sanded to avoid the slightest corrosion and dents. Due to the presence of side recesses in the form of a holder for balls, the carriage moves along the beam. Unlike a round shaft, the collision area is not a point but a line. This creates certain advantages, such as increasing the wear resistance of the rail, increasing the accuracy of the machine, increasing the load capacity, and improving the ability to withstand loads. Accordingly, the minimum backlash or its complete absence is provided. To achieve maximum accuracy in the developed model of the CNC machine, it is proposed to use rail linear guides as shown in Figure 19 [39].

A SM is attached to the aluminum so that its shaft is inside the profile, and the guide is fixed from above by the cut hole from the corner that will fasten the portal with the Y axis. The portal is mounted on the carriage, the operation is symmetrical, and the belt is fastened inside the profile. Figure 20 shows the mounting of the portal with the Y axis.

After that, the design is fastened on the chipboard, the power is supplied, and hence the laser is fastened on a portal carriage. Figure 21 shows a side view.

All electronic parts are placed in the rear and covered with a piece of chipboard resulting in building the body as shown in Figure 22, which shows the model of the assembled CNC laser machine.
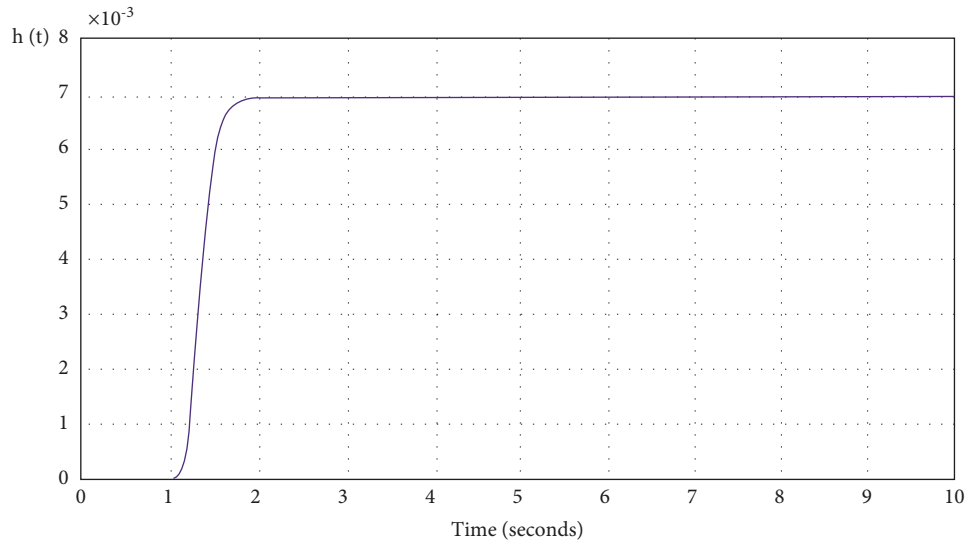
FIGURE 17: Transition process of CNC control system with optimized parameters UNDER the regulator.
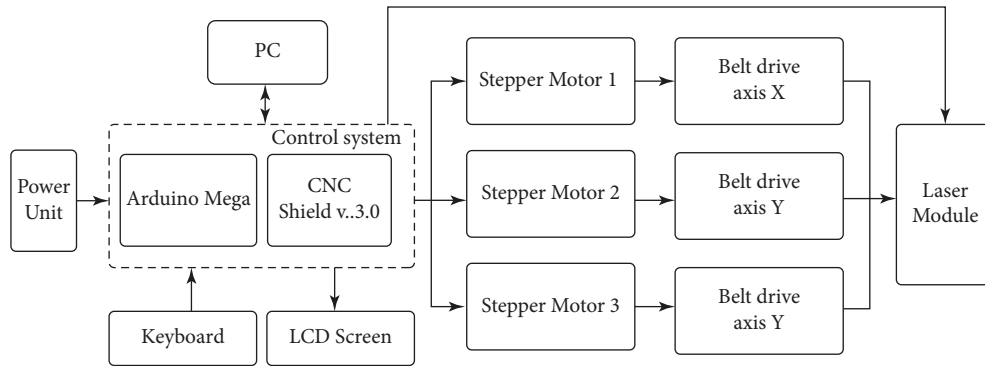
FIGURE 18: Block diagram of a CNC laser machine.
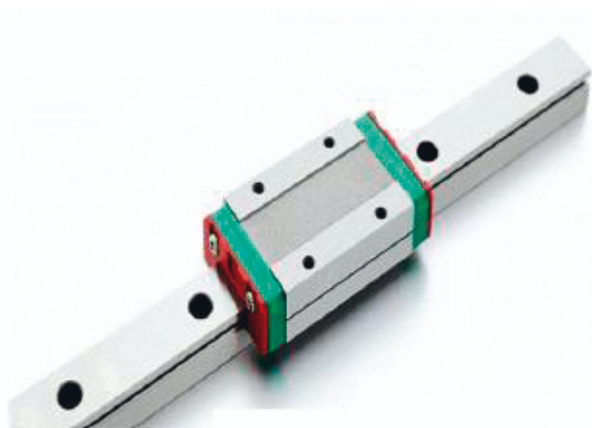
FIGURE 19: Rail linear guide [39].
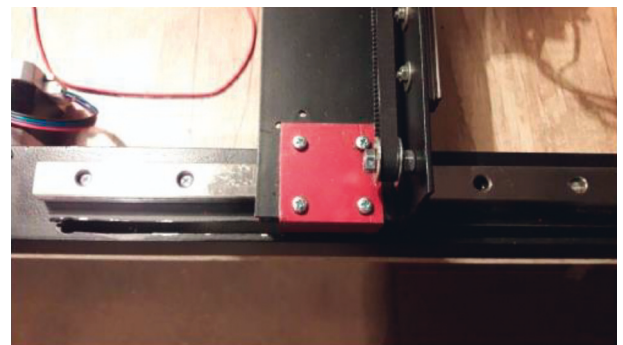
FIGURE 20: Attaching the portal to the Y axis.

the machine with CNC, getting rid of missing SM steps, and reducing the heat of the SM or driver.

To calculate the voltage on the A4988 driver board, the following formula is used [40]:

$$V_{\text{ref}} = I_{\max} \times 8 \times R_s, \qquad (17)$$

where V_{ref} is the voltage on the driver board (v), $I_{\max}$ is the rated motor current (A), and R_s is the resistance of the resistor on the board A4988 ($0.1 \, \Omega$).

After assembling the prototype of the layout, the voltage on the driver board A4988 is adjusted for CD, and finally, the value of the step (Step) of the belt drive is calculated.

The correct current setting for the SM affects the following characteristics: noise reduction from the operation of

FIGURE 21: View of the CNC machine from the side.
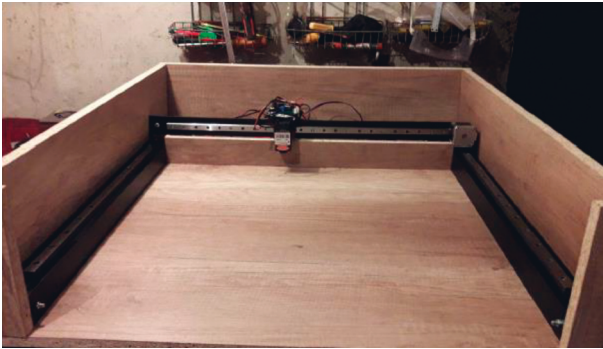
FIGURE 22: The assembled layout prototype of the Laser machine with CNC.

FIGURE 23: The variable resistor on the A4988 board.

TABLE 2: Step settings on the A4988 driver.

MS0	MS1	MS2	Step resolution
Low	Low	Low	Full step
High	Low	Low	1/2
Low	High	Low	1/4
High	High	Low	1/8
High	High	High	1/16

Due to the fact that the operating current of the motor is equal to 70% of the holding current, the obtained value of 1.36 V (from cutout 15) must be multiplied by 0.7. Otherwise, the temperature of SM in the hold mode increases resulting in heating the SM. Accordingly, equation (17) is modified as shown in the following equation:

$$V_{ref} = 1.36 \times 0.7 = 0.952. \tag{18}$$

The result of 0.952 V should be adjusted on the board by adjusting the variable resistor shown in Figure 23.

To "fine tune" the A4988 driver, you need a 5 V voltage source (Arduino Nano controller). You will also need a voltmeter and a small screwdriver. First, we connect the common wire of the Arduino Nano to the common driver and +5 V to the power output of the driver logic called VDD. We power the Arduino Nano and use a voltmeter to measure the voltage between the common wire and the trimmer motor (in the green circle in Figure 23). Next, twisting the construction resistor with a screwdriver, we achieve the calculated voltage of 0.952 V on the voltmeter. The specifics of setting the A4988 driver is that during the setting, it is necessary to carry out without contact of the hand to the metal part of the screwdriver since the hand can act on the measured electrical circuit. This, accordingly, can introduce errors into the settings and, consequently, into the quality of the image obtained by laser engraving. If the voltage does not change when the resistor slider rotates, then the A4988 driver chip is out of order.

To calculate the value of the step (Step) for the belt drive, it is required at first to determine the number of the needed steps to enable the engine's shaft to generate one full revolution, i.e., 360°, so, to obtain 200 steps for example, the single step adjusted to rotate the shaft by 1.8° according to the following formula:

$$N_d = \frac{360}{1.8} = 200, \tag{19}$$

where N_d is the number of the steps needed to obtain one full rotation for the shaft.

Therefore, it takes 200 steps (pulses) to make a full shaft rotation. Moreover, the A4988 driver enables increasing the number of engine steps by controlling the intermediate steps, which increases the accuracy of the CNC machine by reducing torque. To adjust the step on this driver, applying the voltage to certain foams according to Table 2 is required.

Choose the largest division of step 1/16 and install jumpers on the expansion board CNC Shield v3.0 where the A4988 driver board is located.

For example, calculating the number of steps at the full rotation of the motor shaft with the division of step 1/16 can be obtained according to equations (19) and (20):

$$N_d = \frac{360}{1.8 \times 1.6} = 3200. \tag{20}$$

Based on the results of the calculation, it is possible to determine the value of the step for the belt drive according to the formula as follows:

$$N = \frac{(N_d/n)}{N_{step}}, \tag{21}$$

where N is the number of the step for the belt drive, n is the number of gear teeth (GT2 toothed belt with 2 mm pitch installed), and N_{step} is the gear pitch frequency. For the developed layout, the value of the step for the belt drive is 80 ($N = 80$).

3.2. Firmware and Configuration of the CNC SHIELD V3.0 Workbench Software.

Based on the selected hardware modules (Figures 3–5) of the developed laser workbench with CNC and block diagrams (Figure 18), the next step is to develop an information model of software relationships. Figure 24 shows the aggregation of the information model of the relationship between the software and the hardware modules of the workbench with CNC.

In the first step, download Marlin 2.0 Firmware as the most optimized for CNC machines [41]. Open the file X:\∗∗∗\Marlin-2.0.x\Marlin.ino in the Arduino IDE. Reconfigure the configuration.h file. For work with the developed laser machine with CNC. First of all, adjusting the speed of the Serial Port to work with the Arduino Mega 2560-M16U is required, which does not need to change:

 # define BAUDRATE 250000 на *# define BAUDRATE 115200*.

Next, the board to which the laser control module is connected is required to be specified, which depends on the used foam. In the proposed work, the following was specified:

```
# ifndef MOTHERBOARD
# define MOTHERBOARD BOARD_RAMPS_14_EFB
# endif
```

The next step is to specify the used laser, where it is not necessary to specify the same settings:

```
# define EXTRUDERS 1
```

After that, it is necessary to adjust the operation of thermistors used to measure the table and extruder. As the proposed work is for a laser engraver, the firmware must indicate their absence (Appendix A).

Adjustment of the PID controller is carried out in the following parts of the firmware, where it is needed to specify the calculated values:

```
# define DEFAULT_bedKp 79.6
# define DEFAULT_bedKi 87.32
# define DEFAULT_bedKd-4
```

Next, it is needed to spend the mood of the limbs, in the form of maximum and minimum position, and in which direction they should park, which is very important because it depends on the grid of the CNC laser machine.

```
# define USE_XMIN_PLUG
# define USE_YMIN_PLUG
# define USE_ZMIN_PLUG
```

The next step is to specify the logic of the suns, normally open or normally closed. This is done in the following lines of code (Appendix B).

Next, it must be indicated which drivers of SMs are used in the assembled machine:

```
# define X_DRIVER_TYPE A4988
# define Y_DRIVER_TYPE A4988
# define Z_DRIVER_TYPE A4988
```

Now it is necessary to adjust the "steps," i.e., how far the carriage of the laser engraver will travel along the axis. The distance depends on the belt, the driver of the soft engine, the engine itself, or the nut screw. To calculate the steppes according to (18-19).

```
#define DEFAULT_AXIS_STEPS_PER_UNIT {80, 80, 400, 500}
```

Then the speed and acceleration are set, as shown in Appendix C:

The next step is to adjust the direction of SMs rotation, if they work properly, except for the second motor, so it must be inverted:

```
define INVERT_X_DIR false
# define INVERT_Y_DIR true
# define INVERT_Z_DIR false
```

The last step is to specify the direction of parking and the size of the work area, as shown in Appendix D.

After configuring the laser engraver firmware, it is needed to save the Marlin 2.0 Firmware and compile it for testing.

4. Examples of Trial Engraving as Confirmation of the Performance of the Laser Engraver Layout

To verify the accuracy of the decisions made, which hardware and software were made for trial engraving on the developed model of the laser engraver.

In Figures 25 and 26, sample drawings for engraving are illustrated for two different engraving surface materials.

Figure 26 shows the results of engraving of Figure 25 on the engraving surface of plywood 4 mm, where the engraving time is 2 hours 30 minutes. The obtained results of the proposed first example show that the accuracy of the resulting image is regarded as high.

Figure 26 shows the results of engraving, where the high engraving precision is clear. However, despite the high accuracy of engraving for the example in Figure 26, it should be noted that for a more detailed assessment of the quality and accuracy of engraving, it is necessary to conduct a number of test experiments, taking into consideration that such experiments and their full description are beyond the scope of this article, moreover, implementing the deep learning technique can lead to much clearer results such as implementing the deep learning in [42]; however, this is regarded as the future work.

It is important to note that the layout of the laser engraver presented in this work showed its fundamental capabilities for engraving on different media, overall performance, compliance with the theoretical foundations, and practical implementation of the laser engraver proposed in the experimental sample. The conducted experiments confirm that the developed control system based on the PID controller allows high-quality engraving on leather, plywood, and wood. In this aspect, the key is the choice of optimal control parameters for controlling the operation of a laser engraver.

Based on the above, the main goal of this study has been achieved, and an experimental portable laser engraver has
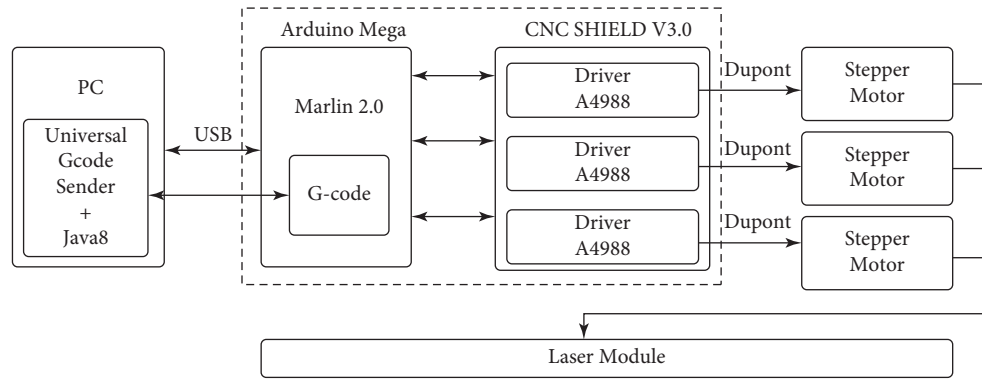
FIGURE 24: Consolidation of the information model of the relationship of the software and the hardware modules of the workbench with CNC.

FIGURE 25: Test image.

FIGURE 26: Results of test image (Figure 25).

been developed and assembled, which has shown its performance and engraving quality.

5. Conclusions

The proposed work in this article is focused on designing and implementing a laser engraving system based on modern industrial Computer Numerical Control (CNC), based machines which shows the great complexity of their application in Armament Repair Shop Set (ARSS) systems. This is due to the large overall dimensions, limited mobility, and the complexity of repairs in the field. The authors made an assumption about the development of small-sized portable laser engravers based on CNC. There is also the possibility of future mobile adaptation to a CNC milling machine by replacing the working body,

which is expected to save a lot of space for other equipment. Based on this solution, the authors set a number of tasks, during the solution of which an analysis was made of the relationship between the input and output parameters of the control object vectors. To comply with these relationships, the authors proposed a block diagram of small-sized portable CNC-based laser engravers. Moreover, the authors developed and mathematically described the control system for selected modules, which applies the theories of automated control. The construction and synthesis of the designed system in MatLab were carried out, and it was proved that the developed system is stable. To ensure the quality and accuracy of laser engraving, the authors optimized the developed system by including a PID controller in it. To check the correctness of the decisions made and the mathematical conclusions, the authors developed an experimental prototype of a small-sized portable CNC-based laser engraver.

In the course of the study, the authors optimized the synthesized control system by using a PID controller in it. By means of mathematical modeling, the indicators of the PID controller were obtained, which accordingly increased the stability of the control system to the occurrence of external vibrations and parasitic factors that could adversely affect the quality of laser engraving.

Finally, the authors applied the proposed model to several actual images, and the obtained results are regarded as high in quality, though just 2 hours and 30 minutes were required to finish the whole laser engraving process in such a small area on different types of surfaces, which practically proves the validity and efficiency of the proposed mode.

The future work is planned to conduct a detailed assessment of the quality and accuracy of engraving based on the developed layout. It is also planned to modify the structures to develop a universal CNC machine with a replaceable working body laser engraver/CNC milling machine.

Appendix

A. Operational Adjustment of the Thermistors

#define TEMP_SENSOR_0 0.ik

```
#define TEMP_SENSOR_1 0.
#define TEMP_SENSOR_2 0.
#define TEMP_SENSOR_3 0.
#define TEMP_SENSOR_4 0.
#define TEMP_SENSOR_5 0.
#define TEMP_SENSOR_6 0.
#define TEMP_SENSOR_7 0.
#define TEMP_SENSOR_BED 0.
#define TEMP_SENSOR_PROBE 0.
#define TEMP_SENSOR_CHAMBER 0.
#define TEMP_SENSOR_COOLER 0.
#define TEMP_SENSOR_REDUNDANT 0.
```

B. Setting the Sums' Logic at Normally Open or Normally Closed

```
#define X_MIN_ENDSTOP_INVERTING false.
#define Y_MIN_ENDSTOP_INVERTING false.
#define Z_MIN_ENDSTOP_INVERTING false.
#define I_MIN_ENDSTOP_INVERTING false.
#define J_MIN_ENDSTOP_INVERTING false.
#define K_MIN_ENDSTOP_INVERTING false.
#define X_MAX_ENDSTOP_INVERTING false.
#define Y_MAX_ENDSTOP_INVERTING false.
#define Z_MAX_ENDSTOP_INVERTING false.
#define I_MAX_ENDSTOP_INVERTING false.
#define J_MAX_ENDSTOP_INVERTING false.
#define K_MAX_ENDSTOP_INVERTING false.
#define     Z_MIN_PROBE_ENDSTOP_INVERTING
false.
```

C. Setting of the Speed and Acceleration

```
#define DEFAULT_MAX_FEEDRATE { 300, 300, 5, 25 }
//#define LIMITED_MAX_FR_EDITING.
#if ENABLED(LIMITED_MAX_FR_EDITING).
#define MAX_FEEDRATE_EDIT_VALUES { 600, 600, 10, 50 }
#endif
#define DEFAULT_MAX_ACCELERATION { 3000, 3000, 100, 10000 }
```

D. Setting the Direction of the Park and the Size of the Work Area

```
#define X_HOME_DIR -1.
#define Y_HOME_DIR -1.
#define Z_HOME_DIR -1.
#define X_BED_SIZE 500.
#define Y_BED_SIZE 500.
#define X_MIN_POS 0.
#define Y_MIN_POS 0.
#define Z_MIN_POS 0.
#define X_MAX_POS X_BED_SIZE.
#define Y_MAX_POS Y_BED_SIZE.
```

Data Availability

We confirm that the data used in this work is available at and ready upon request. Please contact the 4th author (e-mail: viacheslav.liashenko@nure.ua).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] F. Yang and S. Gu, "Industry 4.0, a revolution that requires technology and national strategies," *Complex Intell. Syst*, vol. 7, no. 3, pp. 1311–1325, 2021.

[2] D. Guo, M. Li, Z. Lyu et al., "Synchroperation in industry 4.0 manufacturing," *International Journal of Production Economics*, vol. 238, Article ID 108171, 2021.

[3] M. Ammar, A. Haleem, M. Javaid, S. Bahl, and A. S. Verma, "Implementing Industry 4.0 technologies in self-healing materials and digitally managing the quality of manufacturing," *Materials Today Proceedings*, vol. 52, pp. 2285–2294, 2022.

[4] S. K. Mustafa, V. Yevsieiev, I. Nevliudov, V. Lyashenko, A. R. Alharbi, and W. Rajeh, "HMI development automation with GUI elements for object-Oriented programming Languages implementation," *International Journal of Engineering Trends and Technology*, vol. 70, no. 1, pp. 139–145, 2022.

[5] I. Nevliudov, V. Yevsieiev, J. H. Baker, M. Ayaz Ahmad, and V. Lyashenko, "Development of a cyber design modeling declarative Language for cyber physical production systems," *Journal of Mathematical and Computational Science*, vol. 11, pp. 520–542, 2021.

[6] A. Bécue, I. Praça, and J. Gama, "Artificial intelligence, cyber-threats and Industry 4.0: challenges and opportunities," *Artificial Intelligence Review*, vol. 54, no. 5, pp. 3849–3886, 2021.

[7] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, "Artificial intelligence applications for industry 4.0: a Literature-based study," *Journal of Industrial Integration and Management*, vol. 07, no. 01, pp. 83–111, 2021.

[8] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, "Significant applications of Big data in industry 4.0," *Journal of Industrial Integration and Management*, vol. 06, no. 04, pp. 429–447, 2021.

[9] R. Bonnard, M. D. S. Arantes, R. Lorbieski, K. M. M. Vieira, and M. C. Nunes, "Big data/analytics platform for Industry 4.0 implementation in advanced manufacturing context," *International Journal of Advanced Manufacturing Technology*, vol. 117, no. 5-6, pp. 1959–1973, 2021.

[10] P. K. Malik, R. Sharma, R. Singh et al., "Industrial Internet of Things and its applications in industry 4.0: state of the Art," *Computer Communications*, vol. 166, pp. 125–139, 2021.

[11] A. Kumar and D. K. Sharma, "An optimized Multilayer Outlier Detection for Internet of Things (IoT) Network as industry 4.0 automation and data Exchange," in *International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, Vol 1166*, D. Gupta, A. Khanna, S. Bhattacharyya, A. E. Hassanien, S. Anand, and A. Jaiswal, Eds., Springer, Singapore, 2021.

[12] P. Radanliev, D. De Roure, R. Nicolescu, M. Huth, and O. Santos, "Artificial intelligence and the Internet of Things in

industry 4.0," *CCF Trans. Pervasive Comp. Interact*, vol. 3, pp. 329–338, 2021.

[13] H. T. S. Alrikabi and N. Ali Jasim, "Design and implementation of Smart city applications based on the Internet of Things," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 15, no. 13, pp. 4–15, 2021.

[14] S. Navulur, A. Sastry, and M. Giri Prasad, "Agricultural management through Wireless sensors and Internet of Things," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 6, pp. 3492–3499, 2017.

[15] A. Hamrol, S. Zerbst, M. Bozek, M. Grabowska, and M. Weber, "Analysis of the conditions for Effective Use of Numerically controlled machine tools," in *Advances in Manufacturing. Lecture Notes in Mechanical Engineering*, A. Hamrol, O. Ciszak, S. Legutko, and M. Jurczyk, Eds., Springer, Cham, 2018.

[16] M. Mikhov and M. Zhilevski, "Study of Energy efficiency in a class of CNC machine tools," in *Proceedings of the 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, June 2021.

[17] J. Zinn, B. Vogel-Heuser, and M. Gruber, "Fault-tolerant control of programmable logic controller-based production systems with deep Reinforcement learning," *Journal of Mechanical Design*, vol. 143, no. 7, Article ID 072004, 2021.

[18] A. Łukaszewicz, K. Szafran, and J. Józwik, "CAx techniques used in UAV design process," in *Proceedings of the 2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pp. 95–98, IEEE, Pisa, Italy, June 2020.

[19] Nato Logistics, "Handbook," 2012, https://www.nato.int/docu/logi-en/logistics_hndbk_2012-en.pdf.

[20] Army, "ARSS new build partnership successful," 2015, https://www.army.mil/article/153458/arss_new_build_partnership_successful.

[21] Technical Manual, "Record Details," 2022, https://armypubs.army.mil/ProductMaps/PubForm/Details.aspx?PUB_ID=87745.

[22] Thomasnet, "Top USA and International CNC Machinery manufacturers," 2022, https://www.thomasnet.com/articles/top-suppliers/cnc-machinery-manufacturers/.

[23] Fenix, "Development of a CNC milling machine using a laser," 2018, https://fenix.tecnico.ulisboa.pt/downloadFile/1407770020546367/Dissertation.pdf.

[24] R. H. A. Hellsen, G. Z. Angelis, M. J. G. van de Molengraft, A. G. de Jager, and J. J. Kok, "Grey-box modeling of friction: an experimental Case-study," *European Journal of Control*, vol. 6, no. 3, pp. 258–267, 2000.

[25] Microchip, "ATmega2560," 2017, https://www.microchip.com/en-us/product/ATmega2560.

[26] Handsontec, "3-Axis CNC/stepper motor Shield for Arduino," 2021, https://www.handsontec.com/dataspecs/cnc-3axis-shield.pdf.

[27] Robotshop, "Pololu 8-35V 2A single Bipolar stepper motor driver A4988," 2022, https://www.robotshop.com/media/files/pdf/datasheet-1182.pdf.

[28] Arduino, "Arduino MEGA 2560&Genuino MEGA 2560," 2022, https://www.arduino.cc/en/Main/ArduinoBoardMega2560.

[29] Datasheetcafe, "17HS4401 Datasheet – 40mm, stepper motor," 2022, http://www.datasheetcafe.com/17hs4401-datasheet-stepper-motor/.

[30] Lagunatools, "X|Series C02 laser - Owner's manual - Laguna tools," 2020, https://lagunatools.com/wp-content/uploads/2020/02/2019-CO2-Manual-XSL-MCNCLTL_082718.pdf.

[31] Arduino, "Focusable 500mw 808nm Infrared IR laser diode Dot module 12V + TTL + fan Cooling," 2022, https://www.arduino-tech.com/focusable-500mw-808nm-infrared-ir-laser-diode-dot-module-12v-ttl-fan-cooling/.

[32] Amazon, "405MD-500 405nm focusable Dot laser module violet-blue diode engraving Carve +12V adapter + holder," 2021, https://www.amazon.com/405MD-500-Focusable-Violet-Blue-Engraving-Adapter/dp/B07PXPBQT7.

[33] Amazon, "FB04-500 500mW 405nm blue violet laser module 2∗2.54-2P TTL/PWM modulation for engraver," 2021, https://www.amazon.in/SRBI-FB04-500-Modulation-EleksMaker-Engraver/dp/B07NCTVQQ2.

[34] Dennisxl, "FB03-500 500mw 405nm blue violet laser module 2.54-3p ttl/pwm modulation engraver for eleksmaker," 2022, https://www.dennisxl.com/products/fb03-500-500mw-405nm-blue-violet-laser-module-2-54-3p-ttl-pwm-modulation-engraver-for-eleksmaker_1266178.

[35] S. A. Frank, "Control theory Tutorial," in *SpringerBriefs in Applied Sciences and Technology*, p. 112, Springer, Cham, 2018.

[36] F. Golnaraghi and B. C. Kuo, *Automatic Control System*, p. 994, John Wiley&Sons.Inc, 2009.

[37] W. Bolton, *Control Systems Newnes*, p. 192, 2002.

[38] M. Criterion, "Encyclopedia of Mathematics," 2020, http://encyclopediaofmath.org/index.php?title=Mikhailov_criterion&oldid=49921.

[39] Tuli-Shop, "KWVE and TKVD linear guide rail system INA - Four-row," 2022, https://www.tuli-shop.com/kwve-tkvd-ina-linear.

[40] All3dp, "Vref calculator: how to tune Your stepper driver," 2021, https://all3dp.com/2/vref-calculator-tmc2209-tmc2208-a4988/.

[41] Marlin, "Download Marlin," 2022, https://marlinfw.org/meta/download/.

[42] K. Rezaee, G. Jeon, M. R. Khosravi, H. H. Attar, and A. Sabzevari, "Deep learning-based microarray cancer classification and ensemble gene selection approach," *IET Systems Biology*, vol. 16, no. 3-4, pp. 120–131, 2022.

Hindawi

Research Article

Privacy Law Protection Based on the Information Security Assurance Algorithm

Zhanjiang Wang and Qifeng Yue [ORCID]

School of Law, Shanghai University of Finance and Economics, Shanghai 200433, China

Correspondence should be addressed to Qifeng Yue; summer01@163.sufe.edu.cn

With the continuous development of network technology, the production and lifestyle of human beings are also quietly changing. People find that the network makes life faster and more convenient and also makes our life more "transparent." The openness, sharing, and convenience of the network make it easier to infringe on the privacy rights of others, bringing an unprecedented impact on traditional privacy protection. This paper studied information security assurance-related algorithms and built an information security assurance model. Through this model, some privacy violation events existing in the network today were analyzed. Then, a corresponding questionnaire was constructed according to some of the problems of these events. The questionnaire is mainly used for research on people of all age groups, mainly to study their understanding of relevant legal knowledge and their views on the law. A total of 188 valid questionnaires were collected in this survey. Only 37.23% people knew a lot about the law. Most of them learned it from the Internet, and they rarely learned about it from relevant books. About 70% people believe that the law plays a big role in privacy protection, and they will choose legal help if they encounter privacy violations. However, this is inconsistent with some data obtained from the information security assurance model. Through analysis, it can be seen that although they want legal help, they lack the relevant knowledge of the law, and they do not know whether the privacy of individuals has been violated. Therefore, it is necessary to strengthen personal knowledge of privacy protection laws and improve the personal information legal protection and fair use system.

1. Introduction

Infringement of personal information is concealed, and information users are good at covering up their own behavior with the help of various information technologies, which make it difficult to define and obtain evidence of their behavior. At the same time, the absence of relevant legal systems, insufficient government supervision, and vague industry rules make it difficult to achieve effective legal protection. On the other hand, the lack of protection leads to abuse, and excessive protection pressure will hinder the development of the personal information industry. The inherent conflict between protection and use makes this a tricky one. The invasion of personal privacy in the Internet age has become a hidden worry and trouble for many netizens, and the security of Internet privacy has become a

difficult problem. It can be said that the protection and use of personal information is a relatively new problem in the world. Many experts, scholars, and staff from various countries and regions are struggling to be on the front line of solving this problem. Therefore, how to successfully deal with information security issues to ensure the security of development is of great practical significance to many countries. The more common or deadly information security issues become, the more critical it is to have effective security assurance.

In order to protect the privacy of individuals, researchers have conducted research on the legal aspects of privacy protection. Among them, Mohamed and Chen [1] examined the potential relationship between the existence of a national data protection framework and its impact on e-health. They researched how college students share and

use personal information through online media. The findings will be used to discuss and suggest improvements needed for the current iteration of Malaysia's data privacy laws [1]. Zapadka [2] studied the selected aspects of statistical secrecy, determining the extent to which the data collection and storage processes provided by the Public Statistics Act constitute justified legal interference with the constitutional rights and freedoms of natural people [2]. Danilovic [3] discussed Serbian employees' right to privacy at work in the early twenty-first century and its relationship to international regulations and practices. He proposed the establishment of an internal privacy policy with full respect for data protection regulations, which can provide a feasible and reasonable expectation of privacy at work [3]. Baruh et al. [4] investigated privacy concerns and literacy as predictors of online service and social networking site usage, information sharing, and adoption of privacy-preserving measures. However, their method has not been practically applied.

In order to further strengthen the protection of personal privacy, scholars have begun to study information security algorithms to analyze the network. Among them, Beebeejaun [5] studied the challenges posed by current laws, which are based on assessing the legal framework for the rapid growth of fintech [5]. Dan [6] examined the role of information assurance in managing data security by analyzing the formulation and introduction of the General Data Protection Regulation in May 2018 [6]. Saranya et al. [7] analyzed different types of cloud security issues and used a variety of information security assurance algorithms to generate keys and signatures to provide efficient security [7]. Nabi et al. [8] studied the application logic-related component risk, representing "the process of unifying security assurance properties" to deal with logic vulnerabilities in the system [8]. Alexander [9] investigated whether the AHP model can be effectively applied to the prioritization of information assurance defense-in-depth measures. Findings showed that AHP does not affect the relationship between information technology analysts' prioritization of the five defense-depth dependent variables and independent variables such as cost, ease of use, and effectiveness in protecting an organization's devices from attack [9]. However, the practicality of their algorithm is not great.

This paper established a model of an information security assurance system by studying information security assurance algorithms. Using this model to analyze the current network and then combining it with the questionnaire to study a large number of people, it is found that there are a large number of privacy violations on the network. However, these groups have a lack of legal knowledge about privacy protection, which further exacerbates the lack of privacy. Therefore, new strategies need to be developed to protect personal privacy. The innovation of this paper is as follows: based on the original information security assurance algorithm, it built a more comprehensive information security assurance system model and optimized the algorithm of the model.

2. Information Security Assurance Algorithms and Privacy Protection

2.1. Information Security. The concept of information security has been continuously expanded and deepened with the development of information technology. Initial information security mainly refers to communication security. The current concept of information security refers to information security assurance that comprehensively protects information infrastructure, application services, and information content [10, 11]. The connotation of information security has been extended to the confidentiality, integrity, authenticity, controllability of information, availability of information infrastructure, and nonrepudiation of interactive behavior, and it also includes active defense in information operation conditions. The relationship between information security, information assurance, and information confrontation is shown in Figure 1.

The research content of information security includes two themes: data security and network security. Closely related to this research topic are the security and reliability of computer mainframes and network environments. There are various measures taken for the security and reliability of the computer host and network environment, such as virus prevention, network firewall, and access control [12].

Data security means that the user's data information is not destroyed, the read and write can be consistent, and unauthorized access can be rejected. It is the most fundamental research content of information security. There are many measures to ensure data security, including physical measures, such as strict control of important computer rooms, strict management of hosts, and data backup. It also includes technical measures, such as data consistency check, data redundancy, and network security. In the network environment, network security is an important fundamental measure to ensure data security.

Network security refers to the use of network management control methods and technical measures to ensure the confidentiality, integrity, and availability of data and information in the network environment [13]. The main goal of network security is to ensure that information transmitted over the network and data on network hosts are not added, altered, lost, or illegally read, and host systems in the network are not illegally exploited. These levels are divided into physical security, host system security, and network security detection.

The information security model is an important basis for describing and characterizing the system and information security features, and it is also the foundation of security thinking and technology. The highest-level security system requires formal design and correctness proof [14]. Information assurance is an interdisciplinary subject involving many fields, including not only information security technology but also security management technology, the technical framework of security protection, and the life cycle process of information assurance. Figure 2 shows the "section" of the extended McCumber model in the time dimension. The multidimensional model fully describes the
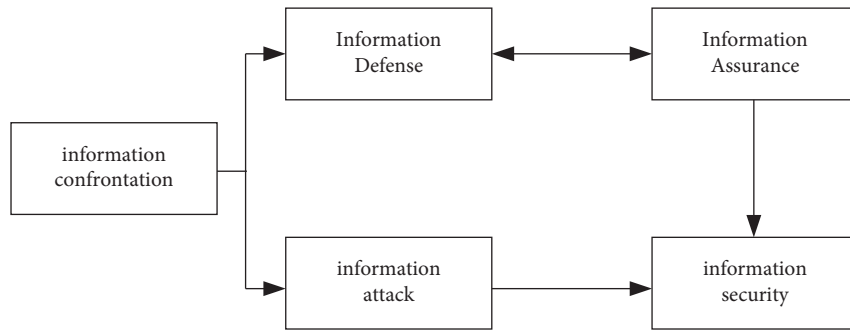
FIGURE 1: Relationship diagram of information security, information assurance, and information confrontation.
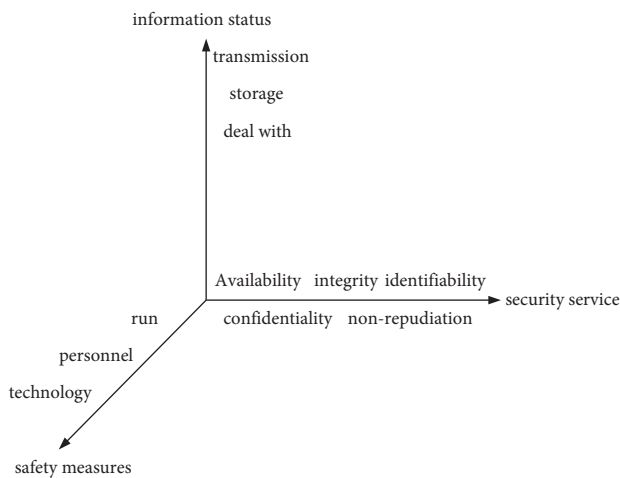
FIGURE 2: Time section of the information assurance model.

FIGURE 3: Basic elements of the network information security system.

various elements contained in information assurance in detail and also reflects the connections and relationships between various elements. The four dimensions included in the information assurance model are information states, security services, security countermeasures, and time.

In order to realize the security of network information, users should establish the capabilities of information collection, protection, and coordination in the information domain. Then, based on this ability, they put themselves in a more favorable position relative to the attacker's information. In this way, the ability to establish, generate, and share high-quality attitude perceptions, jointly understand defense intentions, and self-synchronization of defense processes in the cognitive domain is further realized. As a state of network information security, the network should have the following three specific capabilities: (1) immunity to external damage; (2) to ensure the normal and effective performance of its functions and to maintain its openness and scalability; (3) to protect its integrity and stability [15]. Network information security should include three elements, as shown in Figure 3.

Based on the above analysis, it can be seen that an effective and complete network information security system should have three mechanisms: protection, detection, and response. Based on this, a complete network information security assurance system should include four systems:

protection capability, monitoring capability, management mechanism, and emergency mechanism, as shown in Figure 4.

An intrusion is any attempt to compromise the integrity, confidentiality, and availability of computer system resources or to override the security mechanisms of a computer or network. Intrusion behaviors include the following two aspects: (1) maliciously obtaining system control rights beyond the legal scope; (2) behaviors that cause certain harm to the network system through system vulnerabilities. Intrusion detection is a technique used to detect intrusions in computer networks that violate security policies [16]. By collecting and analyzing key information in the network and host systems in real time, it timely detects and reports the abuse of resources by illegal users and legitimate users in the system and makes appropriate responses. The schematic diagram of intrusion detection and its relationship with other technologies are shown in Figure 5.

From Figure 5(a), it can be seen that the intrusion detection is placed behind the firewall, and its monitoring or
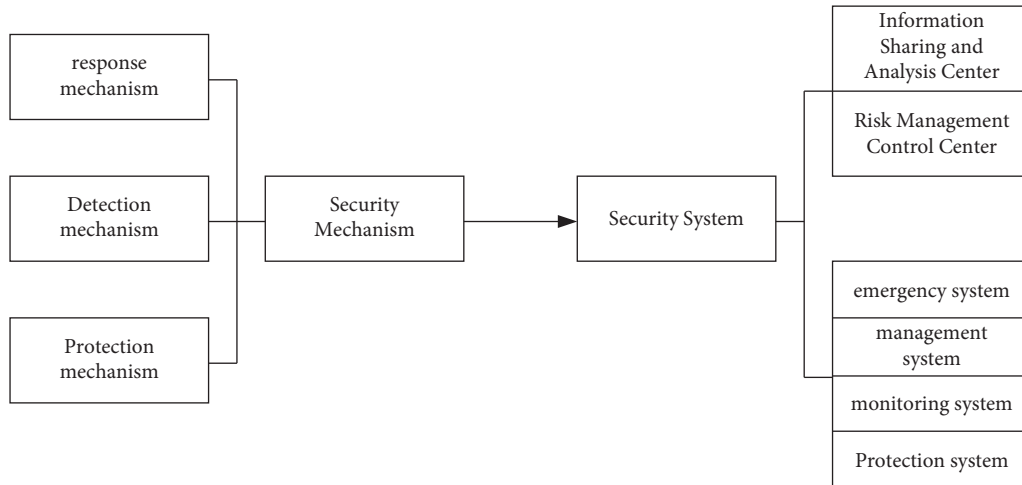
FIGURE 4: Composition model of the network information security assurance system.
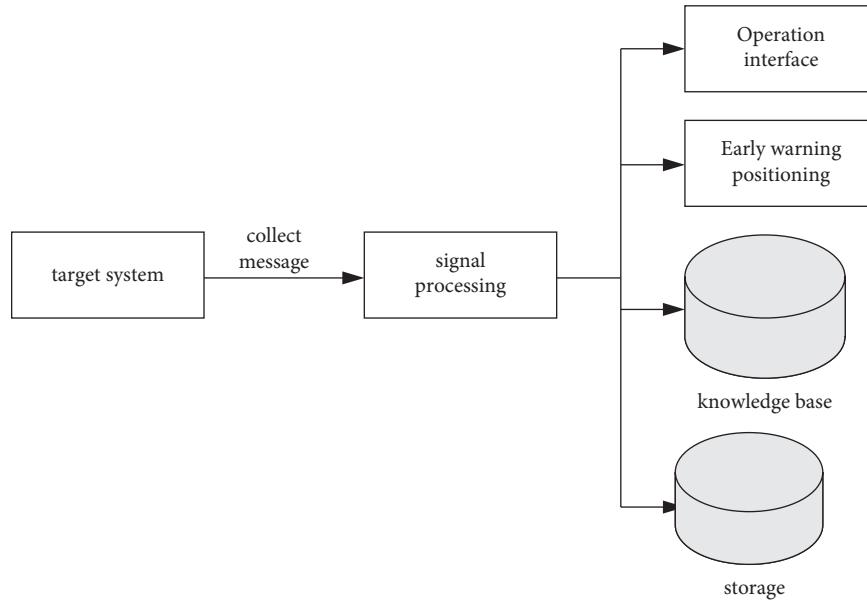
(a)
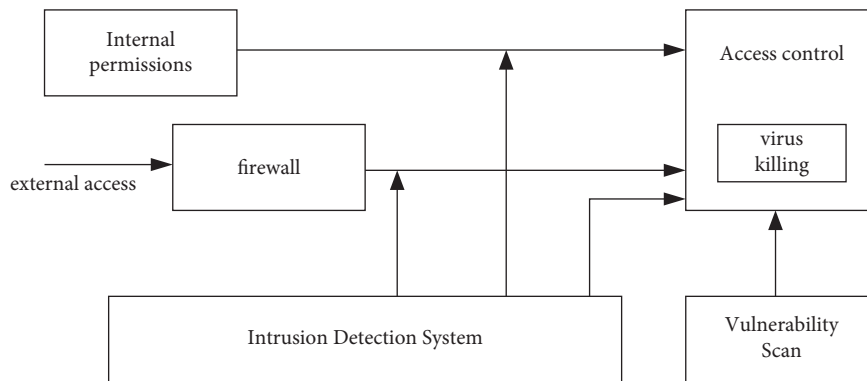
(b)

FIGURE 5: Intrusion Detection. (a) Principle of the intrusion detection system. (b) The relationship between intrusion detection and other security technologies.

information processing of the network generally does not affect the performance of the system. It provides real-time protection against internal and external attacks and more. It can be seen from Figure 5(b) that in the overall security system of a network, intrusion detection systems, firewalls, vulnerability scanning, and antivirus software work together to describe the role of IDS in the network security system [17, 18]. The main role and function of IDS should be reflected in the following points: (1) subjecting the activities of users and systems to its monitoring and analysis; (2) configuration and weaknesses of audit systems; (3) assessing the integrity of critical system data files.

At present, the most widely used intrusion detection systems are mainly expert systems and statistical analysis techniques. The expert system is currently the most widely used technology in abuse detection systems. It uses domain-expert knowledge and reasoning mechanisms and has certain intelligence. It is mainly used for the abuse detection system. The main method of this technology is to set the domain expert knowledge as an intrusion detection rule and judge the legitimacy of the behavior by detecting whether the user's behavior is consistent with the detection rule.

A statistical survey is one of the earliest technologies used in intrusion detection technology, and many commercial versions of intrusion detection products still use this technology [19]. The system assumes that the user's behavioral habits are regular and then builds an auditor of the system by analyzing and modelling the user's historical behavior or collecting early evidence of the system. The auditor judges the intrusion by checking whether the user's behavior conforms to the normal behavior model. Their schematics are shown in Figure 6.

Figure 6(a) shows the schematic diagram of the intrusion detection of the expert system. It can be seen from the figure that the key technologies in the expert system are as follows: (1) The detection knowledge base stores the knowledge of intrusion detection rules. It is operated by experts in the security field and can add detection rules at any time. (2) The inference engine simulates the reasoning ability of the human brain and judges whether the network behavior is legal or not according to the intrusion rules and the rules of the comprehensive knowledge base. (3) The comprehensive knowledge base stores data of various networks and hosts, intermediate inference results, and final results. The system has a high detection efficiency for known intrusion behaviors but does not have the ability to detect unknown intrusion behaviors. The acquisition of intrusion detection rules requires human intervention, and the ability to automatically acquire knowledge is low.

Figure 6(b) shows the schematic diagram of the intrusion detection system based on statistical analysis. It can be seen from the figure that the detection method of the detection system uses some statistical variables to describe the behavior of users or systems. The main functions of the system include selecting appropriate statistics, statistical data analysis and intrusion judgment, and adaptive update of statistical models. Intrusion detection systems based on

statistical analysis have a certain learning ability and adjust the model data of the auditor by continuously analyzing the historical behavior of users or systems. However, it also has a fatal disadvantage that illegal users can train intrusion detectors through a large number of abnormal behaviors to achieve "legitimate" purposes. In addition, the real-time performance of the system is poor.

2.2. Information Security Assurance Algorithm. By means of information security assurance, the risk of information and information systems can be reduced. Considering the differences in the degree of informatization between different countries and the gap in information technology, this paper puts forward the idea of building an information security assurance system based on risk management, highlighting the importance of risk management in an information security system. The information security assurance model is shown in Figure 7.

Through the risk analysis, the security threats existing in the information environment and the vulnerability analysis of the system are given. According to the context of information security assurance and the system vulnerability obtained by the risk analysis, appropriate security control measures are selected. At the same time, the cost of safeguard measures and the value of information assets are analyzed in order to obtain the optimal risk management scheme [20].

The BP neural network is the most widely used network at present, which is usually composed of an input layer, several hidden layers, and an output layer. The BP algorithm is a supervised learning algorithm. Its main idea is to use gradient search technology for known learning sample pairs, in order to minimize the mean square error between the actual output value of the network and the expected output value. If the expected output cannot be obtained in the output layer, it is transferred to backpropagation, and the error of the output signal is returned along the original connection path. The weight of each layer of neurons is modified to minimize the error. The basic evaluation process is as follows:

Network state initialization: assigning a random number between the connection weight w_{ij}, v_{jt} and the threshold θ_j, γ_t endow and $(-1, 1)$.

Entering the first learning sample pair.

Calculating the input u_j and output h_j of each neuron in the middle layer, namely,

$$
\begin{aligned}
u_j &= \sum_{i=1}^n w_{ij} x_i - \theta_j, \\
h_j &= f\left(u_j\right).
\end{aligned}
\tag{1}
$$

Among them, f adopts the sigmoid function, i.e.,

$$
f(u) = \frac{1}{\exp(-u) + 1}.
\tag{2}
$$

Calculating the input l_t and output y_t of each neuron in the output layer, namely,

(a)

(b)
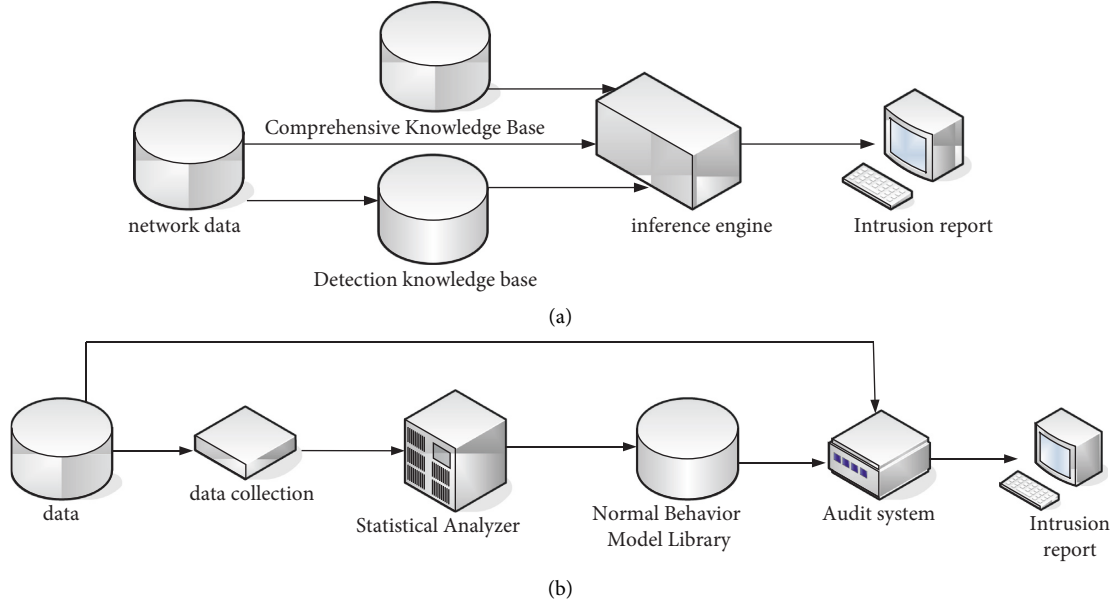
FIGURE 6: Intrusion detection system. (a) The schematic diagram of expert system intrusion detection. (b) Principle of the intrusion detection system based on statistical analysis.
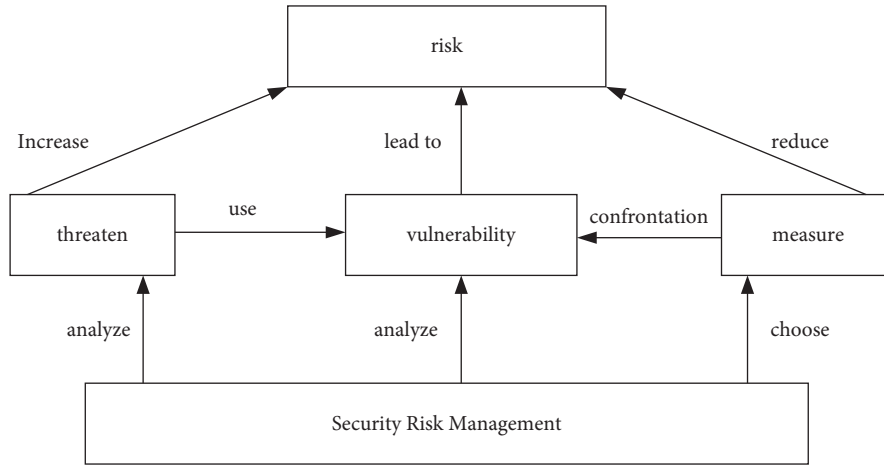
FIGURE 7: Information security assurance model based on risk management.

$$l_t = \sum v_{jt} h_j - \gamma_t,$$

$$y_t = f(l_t) = \frac{1}{\exp(l_t) + 1}. \tag{3}$$

Calculating the weight error δ_t connected to the output layer unit t, which is as follows:

$$\delta_t = (c_t - y_t) y_t (1 - y_t). \tag{4}$$

Among them, c_t in the above formula is the expected value of the sample.

Calculating the weight error σ_t connected to the intermediate layer unit j, as shown in the following equation:

$$\sigma_t = \sum_{t=1}^{q} \delta_t v_{jt} h_j (1 - h_j). \tag{5}$$

Updating the connection weight v_{jt} and the threshold γ_t, i.e.,

$$v_{jt}(N+1) = v_{jt}(N) + \alpha \sigma_t h_j,$$

$$\gamma_t(N+1) = \gamma_t(N) - \beta \delta_t. \tag{6}$$

Updating the connection weight w_{ij} and the threshold θ_j, i.e.,

$$w_{ij}(N+1) = w_{ij}(N) + \alpha \sigma_j x_i,$$

$$\theta_j(N+1) = \theta_j(N) - \beta \sigma_j. \tag{7}$$

Enter the next learning sample pair and return to step 3 until all z pattern pairs are trained.

Starting a new round of learning and training until the following conditions are met:

$$\left| \sum_{k=1}^{z} E_k \right| \leq \varepsilon. \tag{8}$$

Among them, ε is the preset accuracy and E_k is the mean square error, i.e.,

$$E_k = \frac{1}{2} \sum_{t=1}^{q} (c_t - y_t)^2. \tag{9}$$

There are more than 100 original data features in the network, which belong to multifeature, high-noise, and nonlinear datasets. If such high-dimensional data are directly used for classification, the computational load of the classifier will be very large. In particular, there will also be a large number of features irrelevant to data classification in the features, which are more likely to reduce the classification accuracy of the classifier. Therefore, before data classification, it is necessary to perform feature selection on the original network data, that is, data dimensionality reduction. The features that are most conducive to data classification are selected from them, thereby reducing the computational complexity of the classifier and improving the accuracy of classification.

The essence of feature selection is a process of optimizing computation, and pattern recognition is still one of the typical solutions. That is to say, it is necessary to find a criterion and express it with a formula, so that the formula reaches the extreme value in the calculation process [21]. In the process of feature evaluation, the evaluation methods are roughly divided into two categories. One is the distance-based separability criterion, that is, the criterion is based on calculating the dispersion degree of the sample in the feature space; the other is based on the probability density distribution as the basis for judgment.

Supposing the feature candidate set $X = [x_1, ..., x_n]$, in its defined n-dimensional feature space, $d = (x_{ik}, x_{jt})$ is used to represent the distance metric between the k-th sample in the i-th class and the l-th sample in the j-th class. Then, the distance metric value can be defined as per the Euclidean formula and other calculations. Considering the computational complexity of the formula, it is converted into the corresponding overall matrix to measure and calculate, and the overall distribution matrix calculation formula is as follows:

The i-th mean vector can be expressed as follows:

$$\overline{X}^{w_i} = \frac{1}{N} \sum_{x \in w_i} X. \tag{10}$$

The mean vector of the sample aggregate can be expressed as follows:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{M} p(w_i) \overline{X}^{w_i}. \tag{11}$$

Type i covariance can be expressed as follows:

$$\sum_i = \frac{1}{N_i - 1} \sum_{X \in w_i} (X - \overline{X}^{w_i})(X - \overline{X}^{w_i})^T. \tag{12}$$

The total covariance of the sample can be expressed as follows:

$$\sum = \frac{1}{N - 1} \sum (X - \overline{X})(X - \overline{X})^T. \tag{13}$$

The intraclass scatter matrix for the i-th class can be expressed as follows:

$$S_i = E \left| (X - \overline{X}^{w_i})(X - \overline{X}^{w_i})^T \right|. \tag{14}$$

The population within the class scatter matrix can be expressed as follows:

$$S_w = \sum_{i=1}^{M} p(w_i) S_i. \tag{15}$$

The population between the class scatter matrix can be expressed as follows:

$$S_B = \sum_{i=1}^{M} p(w_i)(\overline{X}^{w_i} - \overline{X})(\overline{X}^{w_i} - \overline{X})^T. \tag{16}$$

The population scatter matrix can be expressed as follows:

$$S_T = E \left| (X - \overline{X})(X - \overline{X})^T \right|. \tag{17}$$

S_T and S_B, S_w have the following relationship:

$$S_T = (S_B + S_w). \tag{18}$$

The intraclass scatter matrix represents the scatter of each sample point around its mean. The interclass scatter feature represents the distance distribution between classes, which depends on the sample category attributes and division. The overall scatter matrix has nothing to do with the sample division and attributes.

2.3. Privacy Protection. The essence of privacy mainly includes three parts, namely, private space, information, and activities that are not disturbed by the outside world. It also has the premise that is harmless to society and others. The privacy right is the right of citizens to make their own decisions about their personal life and information and has nothing to do with the public interest.

Privacy protection is mainly performed through the method of government legislation to formulate relevant legal regulations. The basic principles and specific systems for the protection of online privacy rights are legally established, and corresponding judicial or administrative remedies are established on this basis. The use of information makes the traditional closed society transition to transparent society. Personal information should not be locked in a "coffer" but should be put into the "big warehouse" of society to give full play to its value. The legal protection of personal information

should not only protect individuals but also affirm the legality of the use of personal information by information users such as governments and commercial organizations. The purpose of legal protection is to seek coordination between the development and utilization of information resources and the protection of personality rights to reasonable use of personal information.

3. Experiments of Privacy Protection Law

3.1. Questionnaire of Privacy Protection Laws. This questionnaire set up a survey on different groups of people on privacy protection laws and conducted a survey on people of different age groups, as shown in Table 1. A total of 200 questionnaires were distributed in this survey, and 188 valid questionnaires were recovered, with a recovery rate of 94%.

It can be seen from Table 1 that in this questionnaire survey on privacy protection legal issues, the majority of people gave positive responses. Among them, 54.26% were teenagers aged 18–25, which were mainly students, followed by workers aged 25–40, accounting for 28.19% of the total questionnaires. The rest of the questionnaires were mainly attributed to minors under the age of 18 and people over 40. Because they mainly had little contact with the Internet, the ratio is relatively small, accounting for 9.57% and 7.98%, respectively.

In response to the concept of privacy and privacy rights, this questionnaire surveyed and found that they mainly understood privacy and privacy rights in a relatively simple way. They simply believed that it was their own personal information, and privacy rights also stayed in personal information. Only a few people have a comprehensive understanding of privacy and privacy rights, and the survey results are shown in Table 2.

From Table 2, it can be seen that only 57 people have learned about privacy and privacy rights, and the other 131 people have not seriously understood privacy and privacy rights. Thirty of the 53 people aged 25 to 40 have seriously understood it. It shows that more than half of people in this age group will learn about privacy protection laws, which are closely related to their work experience. It makes them more concerned about their personal information. Only 21 out of 102 aged 18 to 25have learned about privacy and privacy rights, which may also include students who majored in law. It shows that contemporary teenagers do not pay much attention to privacy concerns, thinking that these have nothing to do with themselves. However, in fact, privacy and privacy rights are closely related to our daily lives. As for minors, they are less concerned about privacy and privacy rights.

In order to further understand people's understanding of privacy protection laws, this questionnaire also raised a number of legal questions. Among them, about the two questions "How did you learn about privacy protection laws?" and "Do you know more or less about privacy laws?," the received questionnaires have been organized, and the questionnaire data are shown in Figure 8.

From Figure 8(a), it can be seen that people in the age group of 18–40 mainly learn about privacy protection laws

TABLE 1: Survey distribution by age groups.

Age	Number of people	Proportion (%)
<18	18	9.57
18 ∼ 25	102	54.26
25 ∼ 40	53	28.19
>40	15	7.98

TABLE 2: Whether people know about privacy and privacy rights.

Age	Yes (number of people)	No (number of people)
<18	1	17
18 ∼ 25	21	81
25 ∼ 40	30	23
>40	5	10

through the Internet, while those who are underage mainly learn about privacy protection laws through schools and books. For people over 40 years old, because they have little contact with schools and little use of the Internet, their information about privacy protection laws mainly comes from books. For one of them from the school, it may be because he is a school teacher. In Figure 8(b), it can be seen that there are significantly more people who know little about privacy protection laws than those who know a lot about privacy protection laws. This shows that contemporary people do not have a clear understanding of the importance of privacy protection laws, and most people have the mentality that privacy violations will not fall on their heads. However, privacy violations happen to us all the time.

In order to understand contemporary views of today's privacy protection laws, this questionnaire asked the two questions "How effective do you think the laws are in protecting privacy today?" and "Would you seek legal aid if your privacy was violated?." The data are shown in Figure 9.

It can be seen from Figure 9(a) that the vast majority of people agree with the current law and believe that the law can protect privacy. A large part of the population will choose to seek legal protection, but some people do not choose to seek legal help, as shown in Figure 9. Although there are not many people, it further illustrates that the current privacy protection laws may not fully protect privacy. There are still some problems, which make some people not want to seek legal help.

In addition, this questionnaire also conducted a questionnaire survey on the lack of legal knowledge among the population. The questionnaire questions were as follows: "Do you have legal popularization in your area?" and "Do you think the impact of legal knowledge on daily life?." The questionnaire data are shown in Figure 10.

From Figure 10(a), it can be seen that most of the people in the area have no legal publicity, which leads them to not understand legal knowledge, especially some people aged 18–40. From Figure 10(b), most people think that legal knowledge has no effect on daily life, while a very small number of people think that legal knowledge has a deep impact on daily life, especially people aged 25–40. Most of this group of people have married and started a business, so they will be more concerned about the daily life of the family.

(a)

(b)
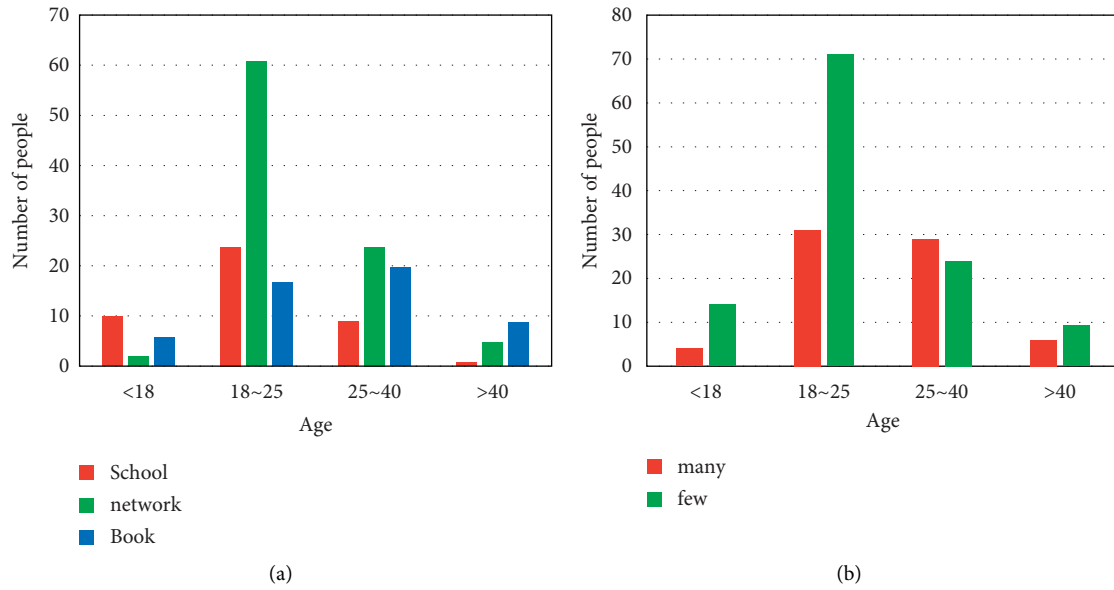
FIGURE 8: Knowledge of privacy protection laws. (a) Access to privacy shield laws. (b) Knowledge of privacy shield laws.

(a)

(b)
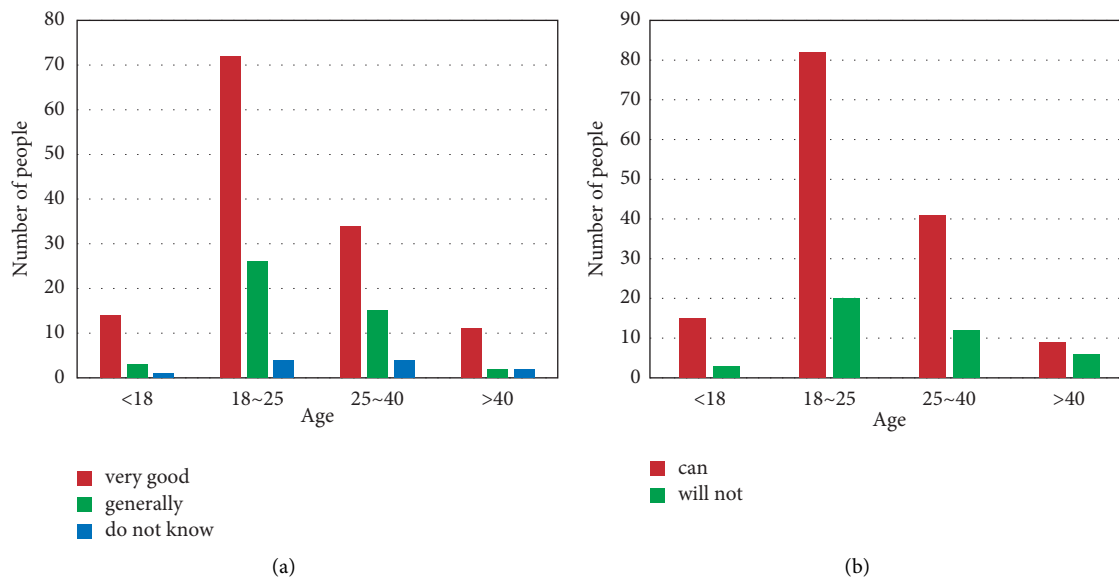
FIGURE 9: Views on privacy protection laws. (a) How effective the law is for privacy protection. (b) Whether they will seek legal aid.

In order to find out whether people are interested in legal knowledge, the questionnaire was surveyed on legal videos. The question was "What would you do if you swiped a short video about the law?," and the data from all the questionnaires were analyzed, and the results are shown in Table 3.

As can be seen from Table 3, the number of people who will watch short videos related to the law is not very large, and most of them will watch for a few seconds or skip them directly. Especially for teenagers in the age group of 18–25, the ratio of being able to watch a complete short video about the law is even less. 28 people will choose to watch it. This is because of the development of technology and the diversity of short videos. Many young people aged 18–25 pursue some funny and celebrity videos and other videos that are not

useful in life, so they will choose to skip short legal videos when they see them. The same is true for other age groups, and the biggest aspect may be the different pursuit of short videos. However, about one-third of people over the age of 40 choose to watch a short video about the law when they swipe it, because this age group has a general awareness of how to make life better.

Combining all data analysis, most people now lack legal knowledge, especially regarding privacy protection laws. Although they know that laws are important, they do not have a detailed understanding of the nature of the relevant laws, which leads them to further ignore the importance of laws. As a result, when their privacy may be violated, they will not choose the law to protect themselves.

(a)

(b)
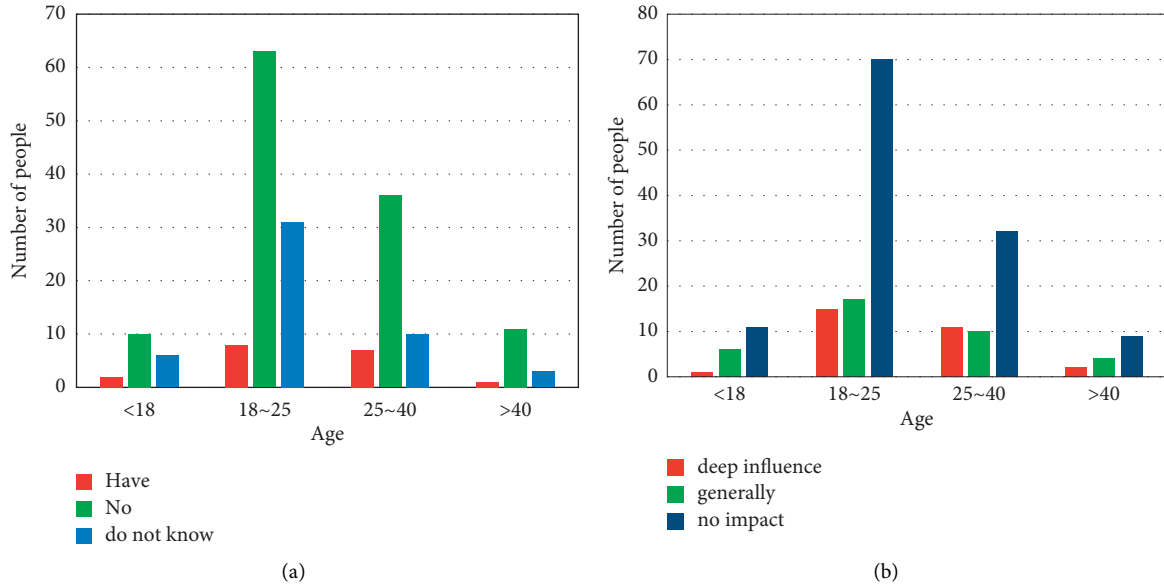
FIGURE 10: Reasons for lack of legal knowledge. (a) Whether there is any publicity about law. (b) Degree of influence of legal knowledge on daily life.

TABLE 3: Practice of brushing short videos on legal aspects.

Age	See it all	Watch it for a few seconds	Skip straight through
<18	2	5	11
18 ∼ 25	28	36	38
25 ∼ 40	11	24	18
>40	5	4	6

3.2. Suggestion. The first point is to improve the legal system. The first is to confirm the right to personal information and to define the subject and object of protection. The struggle for division makes the objects of legal protection have the right to maintain. Second, it is recommended that we distinguish personal information into private information and general information. The unreasonable use of private information is far more harmful to its subjects than that of general information. The protection of the former should be strengthened by the law, and the constraints on the use of the latter should be weakened.

The second point is to build a government supervision system to implement the legal protection of personal information and guide users to use it reasonably. Specifically, through the use of the licensing system, the entry threshold for the personal information industry is raised, and those commercial institutions that do not have the ability to use security measures and regulatory systems are excluded, and the work of the regulatory agency is focused. At the same time, the reward and punishment measures are clearly stipulated to play an exemplary role. Rewards are used to subsidize the efforts of enterprises in the process of reasonable use, whereas punishment is used to prevent abuse of illegal profits.

The third point is to respond to the legal protection of personal information from the perspective of industry autonomy. Through self-management, self-requirement, and self-promotion in the form of industry associations and industry codes, they should do a good job in the self-discipline of the industry and also participate in the balance of protection and use.

The fourth point is from an individual point of view. It is suggested that we improve the liability system and introduce a representative litigation model. The unequal proof capacity and response level should be balanced between information rights and information users to reduce the cost of the former rights of the protection process and increase the amount of damages so as to protect personal information, dispel the speculation of information users, and use personal information rationally.

4. Conclusion

This paper investigated and researched people of all age groups by means of questionnaires, mainly to explore the citizens' understanding of privacy protection laws and other laws. From all the data analysis of the questionnaire, it can be seen that people of all age groups have a serious lack of legal knowledge and little knowledge about privacy protection. As an intangible property, personal information plays an increasingly important role in the development of society. Therefore, it is necessary to strengthen citizens' knowledge of the law so that more citizens can use legal knowledge to protect themselves when they encounter privacy violations. People believe that with the gradual improvement of the country's legislation on privacy protection, the popularization of privacy protection education, and the protection of privacy rights, technological means will change with each passing day.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to sensitivity and data use agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

All authors read and approved the manuscript.

References

[1] A. Mohamed and L. F. Chen, "Data privacy protection: a study on students awareness of personal data privacy protection in an E-health environment," *Journal of Computational and Theoretical Nanoscience*, vol. 23, no. 6, pp. 5299–5303, 2017.

[2] P. Zapadka, "Statistical confidentiality as an element of the protection of the constitutional right to privacy," *Wiadomości Statystyczne The Polish Statistician*, vol. 64, no. 12, pp. 58–67, 2019.

[3] J. Danilovic, "The right to privacy at work," *Anali Pravnog Fakulteta U Beogradu*, vol. 65, no. 2, pp. 162–182, 2017.

[4] L. Baruh, E. Secinti, and Z. Cemalcilar, "Online privacy concerns and privacy management: a meta-analytical review," *Journal of Communication*, vol. 67, no. 1, pp. 26–53, 2017.

[5] A. Beebeejaun, "PRivacy laws in the context of fintech industry in Mauritius: a comparative study," *International Journal of Humanities and Social Science*, vol. 3, no. 4, pp. 2521–2793, 2019.

[6] P. Dan, "The role of information assurance in managing data security," *Database & Network Journal*, vol. 49, no. 1, p. 10, 2019.

[7] J. C. Saranya, V. Usha, and D. S. Alex, "Dynamic data integrity and checkpoint recovery using public auditing in cloud storage," *International Journal of Civil Engineering and Technology*, vol. 8, no. 9, pp. 692–700, 2017.

[8] F. Nabi and M. M. Nabi, "A process of security assurance properties unification for application logic," *International Journal of Electronics and Information Engineering*, vol. 6, no. 1, pp. 40–48, 2017.

[9] R. Alexander, "Using the analytical hierarchy process model in the prioritization of information assurance defense in-depth measures?—a quantitative study," *Journal of Information Security*, vol. 08, no. 03, pp. 166–173, 2017.

[10] F. Nabi, J. Yong, X. Tao, M. S. Malhi, U. Mahmood, and U. Iqbal, "Concepts of safety critical systems unification approach & security assurance process," *Journal of Information Security*, vol. 11, no. 04, pp. 292–303, 2020.

[11] G. Zheng, G. Fang, R. Shankaran et al., "Multiple ECG fiducial points based random binary sequence generation for securing wireless body area networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 655–663, 2017.

[12] D. E. Seidl, P. Jankowski, and K. C. Clarke, "Privacy and false identification risk in geomasking techniques," *Geographical Analysis*, vol. 50, no. 3, pp. 280–297, 2017.

[13] V. V. Derega, "Legal basis of the public policy in the sphere of protection of the right to life for the child at the prenatal stage of development," *Public Administration and Regional Development*, vol. 2018, no. 1, pp. 85–97, 2018.

[14] M. Ozgoli and L. Ariyarathna, "Protection of unpublished works after death: an analysis of developments in Australia and the United States," *European Intellectual Property Review*, vol. 42, no. 5, pp. 312–319, 2020.

[15] J. Sándor, "Whose biobank? Should biobanks serve research interests or the needs for personalized medicine? Analysis of the Hungarian law," *Genetic Testing and Molecular Biomarkers*, vol. 21, no. 3, pp. 140–147, 2017.

[16] H. Schmidt and R. R. Patel, "Should employers Be permitted not to hire smokers? A review of us legal provisions," *International Journal of Health Policy and Management (IJHPM)*, vol. 6, no. 12, pp. 701–706, 2017.

[17] V. Renaudin, A. Dommes, and M. Guilbot, "Engineering, human and legal challenges of navigation systems for personal mobility," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 1, pp. 177–191, 2017.

[18] S. Agrawal, S. Banerjee, and S. Sharma, "Privacy and security of aadhaar: a computer science perspective," *Economic and Political Weekly*, vol. 52, no. 37, pp. 93–102, 2017.

[19] S. R. Muller and M. L. Lind, "Factors in information assurance professionals' intentions to adhere to information security policies," *International Journal of Systems and Software Security and Protection*, vol. 11, no. 1, pp. 17–32, 2020.

[20] J. Al-Muhtadi, M. Qiang, K. Saleem, M. Almusallam, and J. J. P. C. Rodrigues, "Misty clouds–A layered cloud platform for online user anonymity in Social Internet of Things," *Future Generation Computer Systems*, vol. 92, pp. 812–820, 2019.

[21] P. K. Paul and P. S. Aithal, "Information assurance into research level in India: the case of possible research-based degrees," *IRA-international Journal of Management & Social Sciences (ISSN 2455-2267)*, vol. 15, no. 4, pp. 142–149, 2019.

Research Article

Design of Brand Business Model Based on Big Data and Internet of Things Technology Application

Sheng Yao (ID)

Guangdong Fanshi Tuoying Brand Planning Co Ltd, Guangzhou 510640, Guangdong, China

Correspondence should be addressed to Sheng Yao; yaosheng@stu.ahu.edu.cn

Appropriate business models are very important for enterprises. With the advent of the new century, the rapid development and application of technologies such as the Internet of Things and big data have become more and more widely used, which provides great convenience for people's daily life. With the rise and development of the Internet of Things and big data technology, human work and life have become easier and easier, and it has also greatly changed human production and development. Under the fierce market competition, enterprises pay more and more attention to the independent research and development capabilities of their own brand products. With the rapid development of the Internet age, the status of brands in the improvement of business models is also increasing. The brand business model design must not only meet the ever-changing market consumer demand but also conform to the ever-changing fashion trends. The article briefly analyzed the development and application of the Internet of Things and big data technology and optimized the design method of the brand business model by using big data technology in design of the brand business model. On this basis, a decision support method for a brand business model based on a data warehouse was proposed to achieve scientific, efficient, and accurate brand product development. In order to understand consumers' consumption habits and awareness of brands, the article randomly selected 400 people to conduct a survey. It was found that 45% of male respondents preferred commercials with simple and elegant colors; 38% of female respondents preferred commercials with minimalist styles. Modern consumers are mainly pragmatic, and the design of brand business models should be developed in this regard.

1. Introduction

With the advent of the era of big data, human society has entered the digital era. The massive increase and penetration of data and information have driven tremendous changes in people's living and production levels, as well as changes in people's consumption needs. It is driven by the continuous development and progress of high-tech such as information technology, network technology, and social media technology, business trends are rapidly changing and changing, presenting a variety of popular information. The intersection of computer networks, e-commerce, information technology, and other fields has resulted in an ever-increasing amount of massive data. In this case, as soon as the concept of big data appeared, it has received extensive attention from scholars, experts, and managers from all over the world. The reason why big data is "big" is not because it is too numerous, but because it has too many types, such as web page scanning, location information services, online and offline consumption, and so on. Under the background of fierce market competition and economic globalization and competition integration, localization, and homogeneity, more and more enterprises pay attention to corporate image, strengthen the core competitiveness, and improve brand awareness. At the same time, they also recognizes the importance of design to the enterprise.

Economic globalization has made it possible for companies from all countries to participate in world competitions and share the world's advanced development achievements. At this crossroads, which is both an opportunity and a challenge, almost all companies have thought of the importance of brands. Lee and Soon's purpose was to

explore the Apple iPhone jailbreak phenomenon, a new scenario in which a company actively suppresses and discourages loyal consumers from cocreating value and customizing its products. Lee and Soon provided a comprehensive conceptual model to show how different theories of consumer behavior are synthesized in this new phenomenon. Compared to previous work, consumer activists were committed to the brands and products they boycott. Instead of switching to other brands, these jailbreakers and hackers stay loyal to the product and really try to help the brand [1]. Although there is a lot of research on Perceived Brand Globalization (PBG), the concept itself is far from clear and more in-depth research is needed. Mandler aimed to contribute to the global brand literature by proposing a scalable PBG concept and examining the impact of traditional versus global brands in practice. Clearly, the impact of international brands was not all positive. The result provided a clearer perspective on the rise and fall of managers' awareness of brand globalization and called for future research on global brands to be included in all levels beyond the brand, so as to fully grasp this phenomenon, but it's not very practical [2]. Huaman-Ramirez et al. aimed to expand the comprehension of how digital platforms actively impact brand trust by presenting two new variables mediated by brand affection and brand creativity, and by testing the modifying role of consumer e-centrism in these associations. Research had shown that brand globalization was positively correlated with brand sentiment, which in turn was positively correlated with brand trust, which suggested a weaker relationship between brand globalization and brand sentiment for ethnocentric consumers. Brand globalization was positively correlated with brand innovation, which in turn was positively correlated with brand trust, which was a weaker relationship between brand globalization and brand innovation for ethnocentric consumers. This research presented new opportunities for managers concerned with optimizing global brand management. First, the findings suggested that managers were interested in increasing the emotional and emotional aspects of their global brands, making them more trustworthy. Secondly, it was important for brand executives to also highlight the innovative sides of their global brands. Indeed, it was important for practitioners not only to constantly come up with new and exciting propositions for citizens but as well as to keep up with the recent trends in the market. The better managers offered new and helpful resolutions to meet consumer needs, the more consumers would trust these global brands [3]. The results of the study indicated that Sarkar et al. aims to find previous research work on brand cocreation and the role of different stakeholders in brand cocreation and to propose a conceptual tool that can be empirically validated. The results of the study showed that the history of literature presented two main stakeholders, namely, consumers and organizations. The importance of the third important stakeholder, the supplier, had been mostly ignored. However, previous studies had shown that properly managed supplier involvement could reduce the time and cost of product and brand development, and it could provide newer technologies and better quality.

Therefore, the importance of suppliers for the creation of a successful brand together cannot be forgotten. In addition to studying the incentives and results of brand cocreation from the standpoint of consumers and organizations, the aim was to examine the supplier's point of view of the brand cocreation framework. This was the first review of the scholarly literature on brand cocreation from the point of view of three key stakeholders: consumers, organizations, and suppliers [4].

The circular economy (CE) is gaining significance, partly due to the rarity of existing natural assets useable in the context of the environment and changes in collective consumer actions. Cutting-edge technologies such as Big Data and the Internet of Things (IoT) have the possibility to pry the concept of CE from being introduced in organizations and societies and becoming more present in everyday life. Nobre and Tavares conducted a content and social network analysis by using the "R" statistical tool and then made comparisons with a number of current industry initiatives. The results showed that China and the United States were the countries most interested in this area and revealed a context with significant research opportunities [5]. Advances in wireless technologies, computing, and sensing devices, as well as the decreasing cost of these technologies, were driving and accelerating the development of cyber-physical systems. These systems used the IoT model to provide various types of services, such as monitoring, weather monitoring, vehicle traffic management, and controls of production activities. The development and adoption of these systems still face various challenges. On the one hand, the development challenges are primarily related to security, robustness, availability, adequate performance, and optimization of energy savings. On the other hand, the use of these systems generates large amounts of fine-grained data that need to be processed and correlated, commonly through big data analyses to extract useful knowledge that can be used by the software services that control them. Ochoa et al. presented a novel contribution to the design, implementation, and use of these systems as part of a special issue on cyber-physical systems, the Internet of Things, and big data [6]. In addition to the functions of self-monitoring, self-diagnosis, and self-protection, intelligent electrical equipment also has the functions of digitally monitoring electrical parameters, implementing judgments, and intelligently controlling and protecting the monitored objects. Jianhua et al. believed that it was reasonable to analyze and judge the status of in-service power equipment, and to use the method of life cycle management for predictive maintenance, thereby improving the reliability and safety of the entire electrical system. It has become an irreversible trend to use big data-related technologies to process the online monitoring parameters of power equipment, so as to realize the full life cycle management of power equipment. In the future, power equipment manufacturers should provide comprehensive and in-depth services to their customers, that is, become service providers, not just physical products of power equipment. The operating parameters of power equipment are obtained through the Internet of Things, and their life status is analyzed based on big data technology [7].

These studies provided a detailed analysis of brand design and IoT big data technologies. It was undeniable that these studies had greatly contributed to the development of the corresponding fields, and a lot of experience could be learned from methods and data analysis. However, the research on brand business model design based on big data and Internet of Things technology was relatively few and not thorough enough, and it was necessary to fully apply these technologies to the research in this field.

In the research of commercial brand design combined with big data and Internet of Things technology, it was found that the business model of the e-commerce platform has developed by leaps and bounds. Among them, the transaction scale structure of the B2C sales model had increased year by year, accounting for 64% in 2018. In the survey of consumers' consumption habits and brand awareness, it was found that 82% of consumers liked to travel and other outdoor activities and more than 75% of consumers were used to buying goods at special counters in department stores or shopping malls, and about 40% of consumers preferred brand commercials with simple colors and simple style.

2. Design Method of Brand Business Model Based on Big Data and Internet of Things Technology Application

The business model is to combine internal and external factors to form an efficient operation system with unique core competitiveness, so as to achieve the best interests of customers. The rapid development of business models has an important relationship with the advancement of technology, the rise of the Internet, and the rise of e-commerce [8]. The speed and tolerance of consumers accepting new things make many companies adjust their corporate and brand strategies at any time, trying to continuously meet the sales hot spots in the market through efficient design and operation of management activities.

The early marketing methods of enterprises were relatively simple. Generally, enterprises made a series of marketing plans through market research, and finally made decisions and sold products through written data. The whole process was one-way. This one-way information research method had no practical significance to the marketing strategy of enterprises. Enterprises still cannot have a deep understanding of customers, and there was no way to connect the relationship between enterprises and customers, so they could not provide better services for customers [9]. Although some companies further used the method of telephone return visits to customers to improve customers' loyalty to the brand, this kind of telephone consultation service method was very unpopular. This method was not only extremely inefficient but also unsatisfactory [10]. After that, many companies also discovered the limitations of the telephone return visit method, so some companies pushed the promotion activities of the merchants through SMS to attract customers. However, the time for this passive marketing method to benefit was very short. On the one hand,

customers were easily numb or even annoyed by this behavior. On the other hand, because there was no way for enterprises to screen according to the needs of customers, and could not convey effective information to customers who really need it, the effectiveness of the marketing method of SMS push was obvious [11].

With the development of science and technology, the Internet has gradually become known to people. By building official brand websites, companies allow customers to understand corporate and brand information more quickly. Enterprises have established the most direct contact information with customers. Customers can find out about enterprises at any time by searching for enterprises and brands on the Internet platform, and can quickly obtain the contact information of enterprises. Although this method of network construction can let people know about enterprises and brands in a faster and more convenient way, this method of publicity is still single information dissemination, and enterprises are still unable to actively provide timely services to customers in need until the mobile internet appeared [12]. If an enterprise does not have Internet thinking and does not know how to apply the large platform of the Internet, it cannot become bigger and stronger if it is limited to traditional store sales. Under the background of the popularity and history of today's mobile Internet, enterprises not only have to have Internet thinking to reasonably apply network platforms for strategic marketing but also need to follow the current trend of technology—WeChat marketing thinking [13]. As an indispensable communication method in people's daily life, WeChat is used as an important means of information acquisition, and WeChat brings customers together. Through WeChat, a well-known communication platform, the communication method between the enterprise and the customer is close to zero distance like relatives or friends. By following the company's WeChat public account, customers can learn about the company's dynamics and brand culture, and obtain information on the latest research and development products and activities. And enterprises can manage more effectively for different customer groups by establishing a complete customer information database. This not only facilitates timely communication and effective interaction between enterprises and customers but also pushes the latest product and activity information developed by the company to customers. It can also answer all customer questions and receive feedback and suggestions from customers through intelligent replies and artificial real-time [14]. Many commodities, like the previous products, are aimed at primary consumers. They are directly sold to the market without packaging and have no inherent brand image. Most of the target groups are grassroots working people as the main consumers. With the rapid development of the Internet industry and the continuous expansion of the market, the positioning of the original commodities makes them face the bottleneck of sales, which increases the unnecessary waste of most commodities [15]. The current sales situation is gradually declining. In other words, in order to change the current low development situation, it is necessary to change the positioning of the products. It is necessary to ensure the intimacy

of products from nature, and it is necessary to meet the sales concept that products come from pure nature without additives, so as to meet people's emotional and psychological needs, make the product itself have a human touch, and stimulate consumers' desire to buy. It is also necessary to build people's loyalty to the brand [16]. Brands not only represent the main products and services in the exchange process but also reflect consumers' deepening psychological feelings and acceptance of commodity value. However, with the changes in the consumption behavior of consumer groups and the turning of their habits, in the face of the rapid changes in people's consumption behavior, some well-known brands with a long history of precipitation cannot fully meet the needs of all consumers. Therefore, the products of the e-commerce platform are resegmented again, and the target market positioning is refined again. The repositioning of consumer groups makes the brand younger, and establishes an internal connection between the brand and this normal consumer group, as shown in Figure 1 [17].

In today's era, the brand not only represents the image of an enterprise, but also a symbol of the enterprise's management strength and innovation ability. The number of world-famous brands also symbolizes the strength of the country's comprehensive competitiveness. The material composition of the value of the brand business model is mainly reflected in the technical level of the product, the design quality of the product, and the manufacturing quality of the product. One of the key points involved in the brand concept is the establishment of brand image, which is based on the perception subject—consumer usage and perception experience—and then it forms the psychological association to the brand. In terms of expression, it includes two aspects (1) Tangible content (external): the functional content of the brand, which is the functional needs of consumers that the products and services of the brand can satisfy; that is, the external intuitive features of products, logos, packaging, etc., which can be directly contacted and seen by consumers, and the intuitive feeling of brand image on the basis of material [18]; (2) Intangible content (intrinsic): the charming characteristics of the brand, that is, the personality characteristics that the entrepreneur endows the brand and is perceived by consumers. It usually includes the brand's cultural connotation, entrepreneurial spirit, and other internal characteristics, which are the key factors to determine the success of the brand [19], as shown in Figure 2.

The derivation of brand value is nothing more than accompanying the growth of customer demand. The promotion of brand value and the creation of customers are inseparable and complementary, and they together constitute the cornerstone of the sustainable development of enterprises [20]. Generally speaking, brand value is mainly composed of brand premium, satisfaction, loyalty, perceived quality, leadership, perceived value, brand personality, organizational association, brand recognition, market share, price, distribution index, etc. [21]. Almost most of them are customer value-oriented factors, so a customer value model based on brand equity is used to illustrate it vividly, as shown in Figure 3. And brand maintenance is particularly important as an inherent requirement for corporate image and

large-scale long-term development. Brand maintenance includes standardized management and testing of product quality, operating technology, and related equipment. More importantly, it is necessary to continuously improve and implement corporate culture and promote the overall improvement of corporate management.

Internet of Things (IoT) is a system that connects people to people, things to things, and people to things for information exchange through various proprietary networks, such as communication networks, sensor networks, and industry-specific networks. Currently, with the development of an increasing number of sensing technologies, such as infrared and 4G, IoT is defined as a network where all objects are connected to the Internet for communication through sensing devices. The current IoT architecture is usually considered to contain three layers: sensing layer, network layer, and application layer, as shown in Figure 4.

With the development of science and technology, people have entered the era of "big data," and big data is becoming the core competitiveness of business competition in this era. The generation, development, and application of big data pose challenges to traditional business models and provide opportunities for continuous innovation of business models. From the previous summary and analysis, it can be seen that big data has an important impact on the content of the four elements of business model innovation. The realization of business model innovation or transformation from the perspective of these four elements is an important condition for enterprises to realize business model innovation. Therefore, this paper constructed a model of the impact of big data on business model innovation, as shown in Figure 5.

Along with the speedy progress and application of information technology, computer technology, and Internet technology, human society has rapidly moved into a modern digital era. With the development of the network, the continuous development of the Internet, mobile, social networking, and other fields, the application scope of network information technology has been greatly expanded. As a positive factor, "data" is constantly eroding people's lives, and the explosion of information has led to a large amount of data. The continuous increase in the amount of information has led to qualitative changes in people's thinking and behavior, which also heralds the advent of the era of big data.

It is supposed there are N customers with K brands to choose from, or they can choose not to buy, that is: customer n ($n = 1, 2, \ldots, N$) has $K + 1$ kinds of products to choose from, $k = 0, 1, \ldots, K$, respectively. For the i-th customer facing $K + 1$ choices, it is supposed the utility of choosing k is

$$U_{nk} = X_n' \beta_k + \varepsilon_{nk}, n = 0, 1, \ldots, K. \tag{1}$$

Among them, X_n' is a vector representing the intrinsic features of the customer, β_k is also a vector representing the parameters for the intrinsic features of the correspondent customer when the k-th good is selected, and ε_{nk} is a random term that is separate and obeys.

$$\varepsilon_{nk} \sim N(0, 1). \tag{2}$$
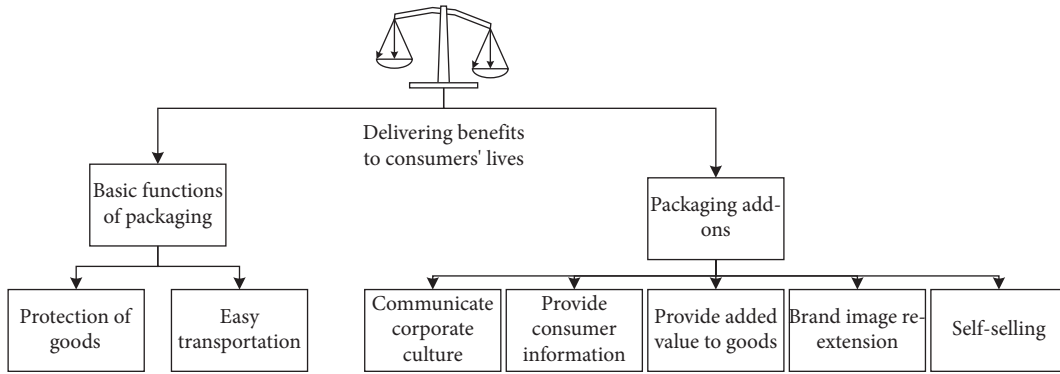
Figure 1: Requirements for brand design in the existing consumer market.
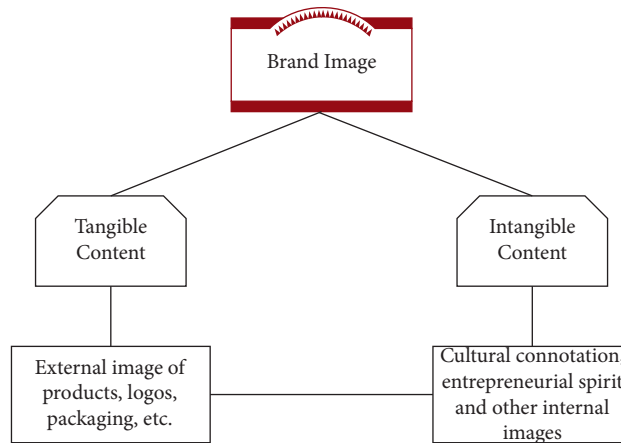
Figure 2: Brand image content.
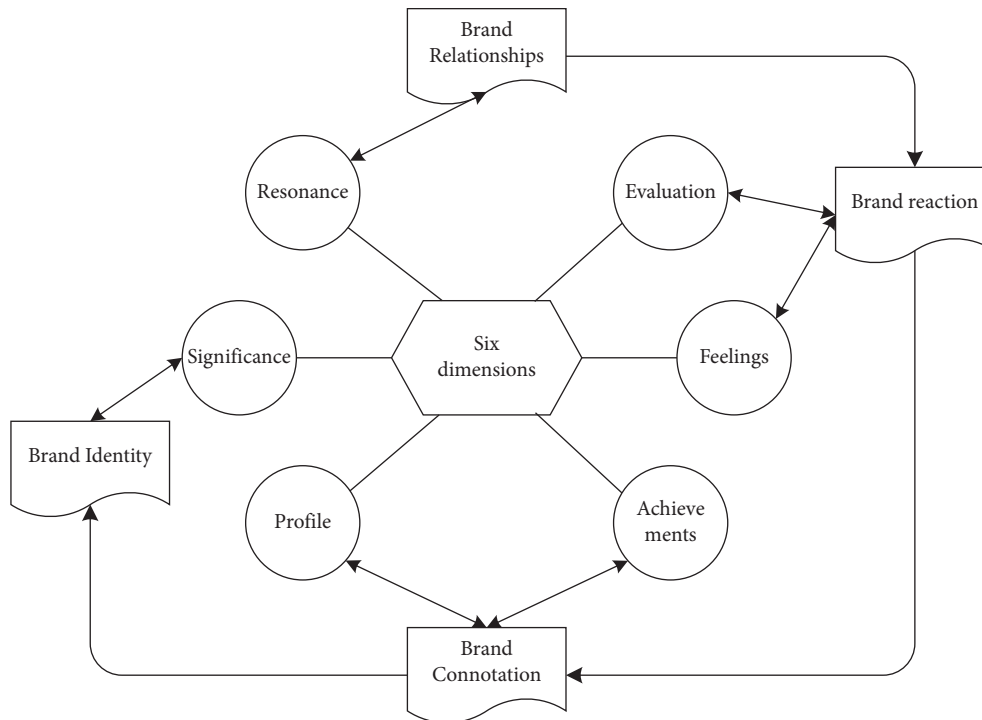
Figure 3: Customer value model based on brand equity.
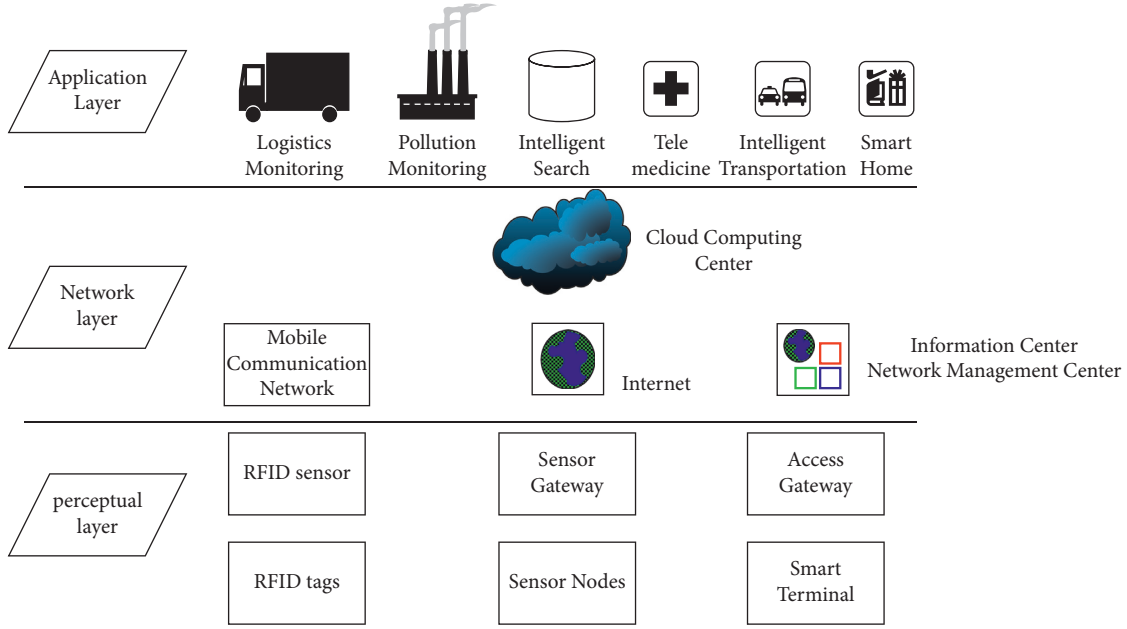
FIGURE 4: Schematic diagram of IoT architecture.
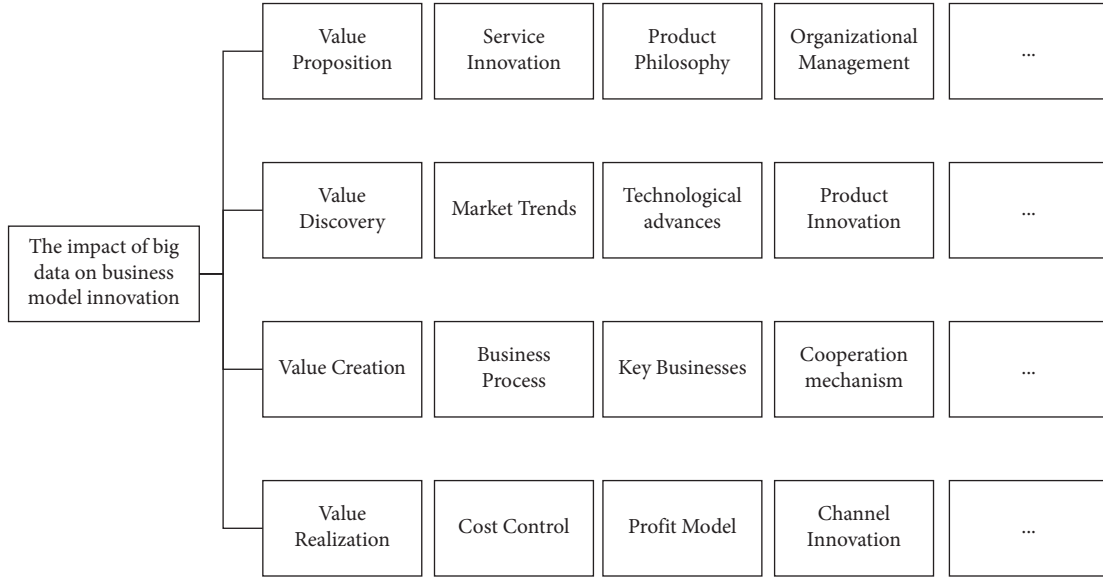
FIGURE 5: Model of the impact of big data on business model innovation.

p_{nk} is to express the possibility of customer n choosing brand k. Then based on the principles of utility maximization, the probability of a customer n choosing brand k can be obtained as follows:

$$p_{nk} = P\left(U_{nk} > U_{nj}, j \neq k\right). \tag{3}$$

The density function for the unobserved utility function is

$$\mathrm{f}\left(\varepsilon_{nk}\right) = \mathrm{e}^{-\varepsilon_{nk}} \mathrm{e}^{-\mathrm{e}^{-\varepsilon_{nk}}}. \tag{4}$$

After integration, the cumulative distribution function is obtained as follows:

$$\mathrm{F}\left(\varepsilon_{nk}\right) = \exp\left(-\mathrm{e}^{-\varepsilon_{nk}}\right). \tag{5}$$

From the random utility part, it can be known that the probability of a customer n choosing option i is

$$\begin{aligned} p_{ni} &= P\left(C_{ni} + \varepsilon_{ni} > C_{nk} + \varepsilon_{nk}, \forall k \neq i\right) \\ &= P\left(\varepsilon_{nj} < \varepsilon_{ni} + C_{ni} - C_{nk}, \forall k \neq i\right). \end{aligned} \tag{6}$$

If ε_{ni} is given, this expression is the cumulative distribution function of every ε_{nk} on $\varepsilon_{ni} + C_{ni} - C_{nj}$, which can be expressed as $\exp\left(\exp\left(\varepsilon_{ni} + C_{ni} - C_{nj}\right)\right)$ by formula (5). Since ε_{nj} is independent, this cumulative distribution is the result of multiplying a single cumulative distribution for all $k \neq i$;

that is, in the presence of random interference, the conditional probability that customer n chooses i is as follows:

$$p_{ni}|\varepsilon_{ni} = \prod_{k \neq i} e^{-e^{-(\varepsilon_{ni}+C_{ni}-C_{nk})}}. \tag{7}$$

Formula (7) is substituted into the distribution to solve p_{ni}

$$p_{ni} = \int (p_{ni} \mid \varepsilon_{ni}) d\varepsilon_{ni}. \tag{8}$$

It can get

$$\begin{aligned} \% \quad p_{ni} &= \int_{\varepsilon_{ni}=-\infty}^{+\infty} \exp\left(-e^{-\varepsilon_{ni}} \sum_k \exp\left(-(C_{ni}-C_{nk})\right)\right) e^{-\varepsilon_{ni}} d_{\varepsilon_{ni}} t \\ &= e^{-\varepsilon_{ni}}. \end{aligned} \tag{9}$$

The formula (9) is substituted to get

$$\begin{aligned} p_{ni} &= \int_{t=\infty}^{0} \exp\left(t \sum_k \exp\left(-(C_{ni}-C_{nk})\right)\right) dt \\ &= \frac{\exp\left(t \sum_k \exp\left(-(C_{ni}-C_{nk})\right)\right)\Big|^{0}}{\sum_k \exp\left(-(C_{ni}-C_{nk})\right)}\Bigg|_{\infty} p_{ni} \\ &= \frac{1}{\sum_k e^{-(C_{ni}-C_{nk})}} = \frac{e^{C_{ni}}}{\sum_k e^{C_{ni}}}. \end{aligned} \tag{10}$$

And from formula (1), it can get

$$p_{ni} = \frac{e^{x_n \prime \beta_i}}{\sum_{k=0}^{K} e^{x_n \prime \beta_i}}. \tag{11}$$

If $\beta_j^* = \beta_j + m$ is defined for any vector m, and then the probabilities defined in formula (11) are recalculated using β_j^* instead of β_j, the result is the same as using β_j, because the terms containing m are eliminated. That is, adding any customer attribute variable m has no effect on the results, so it can be concluded that customer attributes, such as gender and occupation., have no effect on the results. For this purpose, $\beta_0 = 0$ can be normalized. (Because the probabilities sum to 1, only K parameter vectors are needed to determine $K+1$ probabilities) so the probabilities are

$$p_{ni} = \frac{e^{x_n \prime \beta_i}}{1 + \sum_{k=1}^{K} e^{x_n \prime \beta_k}}. \tag{12}$$

By formula (12), the K log odds ratios can be calculated as follows:

$$\ln\left(\frac{p_{ni}}{p_{nk}}\right) = \ln \frac{e^{x_n \prime \beta_i}/1 + \sum_{k=1}^{K} e^{x_n \prime \beta_k}}{e^{x_n \prime \beta_k}/1 + \sum_{k=1}^{K} e^{x_n \prime \beta_K}} = x_n \prime (\beta_i - \beta_j), j \neq i. \tag{13}$$

If the model has $K+1$ categories, it can be expressed as follows:

$$\begin{aligned} \ln\left(\frac{p_{n0}}{p_{nK}}\right) &= x_n \prime (\beta_0 - \beta_K), \\ \ln\left(\frac{p_{n1}}{p_{nK}}\right) &= x_n \prime (\beta_1 - \beta_K), \\ \ln\left(\frac{p_{n(K-1)}}{p_{nK}}\right) &= x_n \prime (\beta_{K-1} - \beta_K). \end{aligned} \tag{14}$$

Among them, K is the reference category, that is, k is the standard to compare with other options.

In product development, the use of big data resources—mainly consumer input and the objective grasp of popular information—this open and innovative way can enhance the specific operation in the design work and reduce the cost of research and development, so as to achieve the ability to respond quickly to consumer and market changes. At the same time, it also plays a key role in realizing the practicability, science, and value of product design.

Through the theory of value sharing, enterprises, customers, and other stakeholders must establish sufficient connections, maintain continuous interaction, and form effective aggregation, so as to promote the overall value creation of the enterprise system. Based on the brand community, it can be divided into three business models as follows: (1) take the enterprise as the center, strengthen the company's functions, and improve the customer's products and services; (2) taking the customer as the center, create customer experience shaping with a special experience for customers; (3) focus on stakeholders, and build a broad set of values for members of society by strengthening the connection of various stakeholders, as shown in Table 1.

3. Experiment Preparation for Brand Business Model Design Based on Big Data and Internet of Things Technology Applications

Successful brands usually have a strong brand identity with concise expressions, diverse forms, and distinct personality traits. A representative brand image will show the essential characteristics of the company and show it in a specific way, allowing customers to have a visual impact, thereby improving the recognition of the product and distinguishing it from other products. As a corporate brand design, there should be distinct differences to facilitate the understanding of customers; at the same time, elements that best reflect the company's characteristics should be incorporated into the design.

In today's increasingly abundant information resources, massive information resources will become an important part of the development of human society. The use of information technology to assist product development has become a new development trend. In the process of information technology guiding product design, obtaining consumers' input and contributions through a large number of consumer information resources is an important means to

TABLE 1: Types of business model innovations based on brand communities.

Name	Objectives	Functions	Community characteristics
Corporate function enhancement	Strengthen the business	Enhance corporate functions	Company-led community result-oriented
Customer experience shaping	Strengthen customers	Focus on a lifestyle theme and unique lifestyle	The natural integration of brand value proposition and lifestyle brand community and corporate business activities are moderately divided by customer-led brand community
Cross-border platform integration	Strengthen stakeholder connections	Stakeholders form connections to generate value	Enterprise has brand appeal
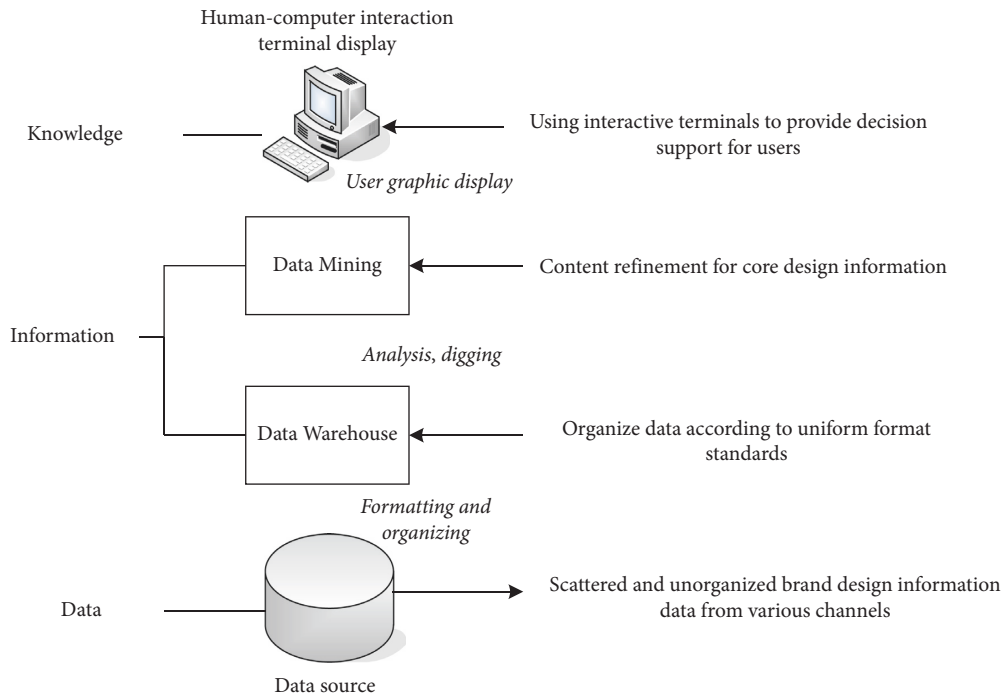
FIGURE 6: Basic process of brand design decision support system based on a data warehouse.

complete product development. Numerous good companies have begun trying to mine and quantify consumer opinions and preferences from traditional data sources, including point-of-sale data and customer feedback. They distill key information to aid product design from new data sources including consumer reviews on social media and third-party platforms, and sensor data describing actual product usage. The decision support method for brand product design based on data warehouse is shown in Figure 6.

At the end of the data collection phase, the company has a complete set of data collection and archiving tools and can either sell the collected data to larger companies that can analyze it, or provide data collection and archiving services to companies without data collection systems. Once data analysis is complete, the company is again fully capable of data extraction and can connect data extraction activities. Once the analysis and statistical reporting are complete, the company can sell the net data mining business.

Once the data application phase is complete, companies are no longer limited to data sales policies. Selling data has

no commercial value for the company's core business, but can adjust and design business around the relationship between data, and effectively implement data-driven business.

The terminal is at the end of the supply chain and is the place where consumers have the closest contact with the brand, so the image of the terminal directly affects the image of the brand in the minds of consumers to a certain extent. And the brand is generally independent, there are franchise chains, there are entrusted companies, and the standard franchise image is uniformly designed. Some brands have high franchise fees and need an optimal location. Classic business districts do not allow returns, and these strict conditions cannot be returned. These strict franchise conditions can control the expansion speed of the brand, and also ensure the relevant service quality of the store to a certain extent. In order to understand consumers' consumption habits and awareness of brands, 400 consumers in area A were randomly surveyed. The basic information of the respondents is shown in Tables 2 and 3.

TABLE 2: Basic information of respondents.

Test items		Number of people	Proportion (%)
	18–25	69	17.25
	26–35	136	34
Age	35–45	156	39
	45–60	19	4.75
	> 60	20	5.0

TABLE 3: Income of respondents.

Test items		Number of people	Proportion (%)
Gender	Male	200	50
	Female	200	40
	2000–3500	50	12.5
	3501–5000	52	13.0
Monthly income range	5001–8000	230	57.5
	8001–12000	46	11.5
	> 12000	22	5.5
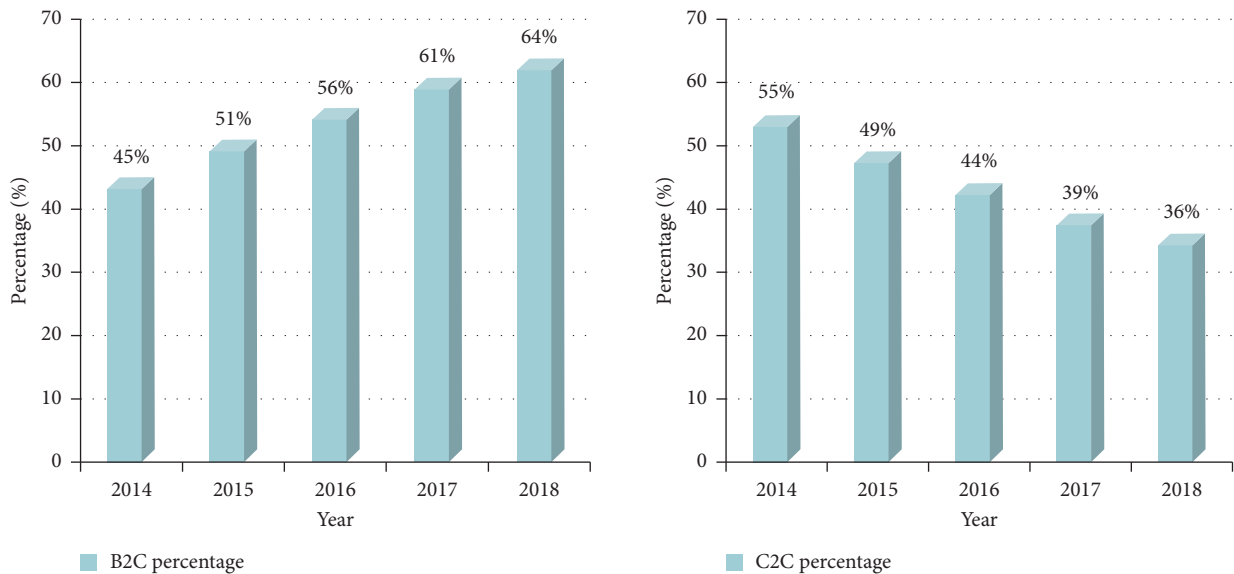
FIGURE 7: 2014–2018 China online shopping market transaction scale structure.

4. Brand Business Model Design Data

In recent years, the development of e-commerce has been very fast, and the types of advertising campaigns will also change at different times, so many sales models will gradually shift from physical stores to e-commerce platforms. In this way, e-commerce can not only expand development in popular industries but also reduce costs for the business itself. This improves the display of products in the sales process, it will be more efficient to face these channels, and the development of e-commerce will be more efficient. Today, when the current monopoly model is changing, in order to meet the needs of customers, future product sales will move in the direction of "digitization, channelization, and platformization." For most consumers, what they like is the traditional handmade products they make or the deep cultural heritage they have. The current e-commerce

platform sales model has become a major sales method, especially since the B2C transaction scale is increasing every year, and by 2018, it has reached 64%, as shown in Figure 7, this sales model has been affected by favored by many young people. If the development model of the e-commerce platform cannot be well grasped, the sales model of professional brick-and-mortar stores will face difficulties and bottlenecks in development.

Figures 8 and 9 are the statistical results of the respondents' brand tendencies.

It can be seen from Figure 8 that the highest proportion of respondents' hobbies is travel, accounting for 82%, followed by music 64%, sports 42%, drama 28%, and art 26%.

As can be seen from Figure 9, 81% of consumers are accustomed to buying things on department store counters, and 78% like to buy things in shopping mall specialty stores. More than 40% of consumers shop through trend
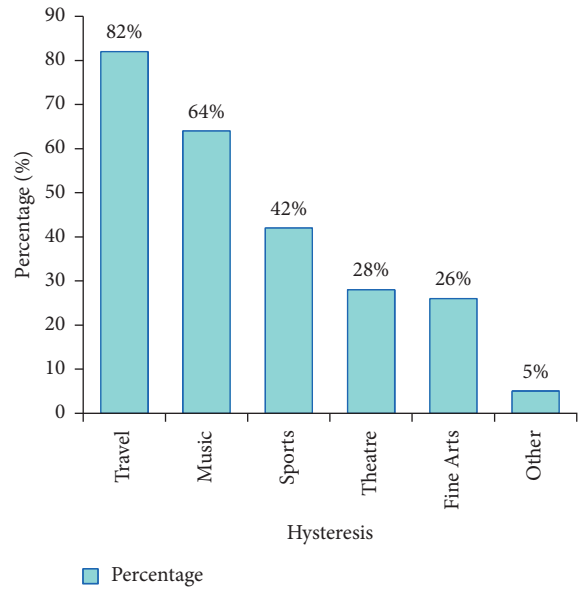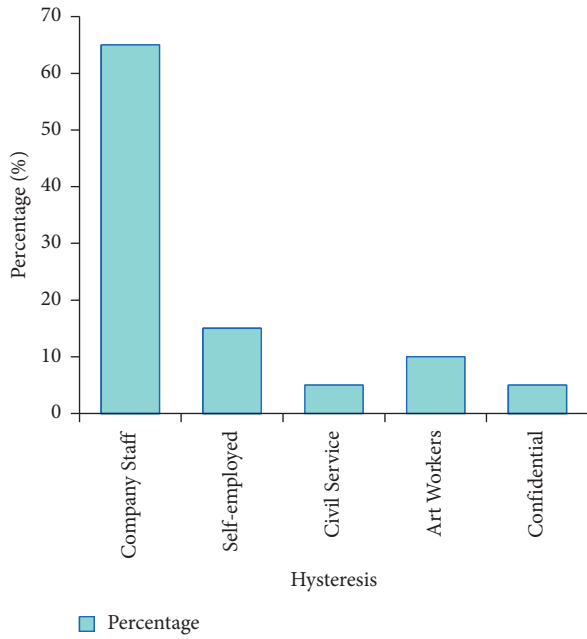
FIGURE 8: Respondents' occupations and hobbies.

information provided by the Internet, fashion magazines,
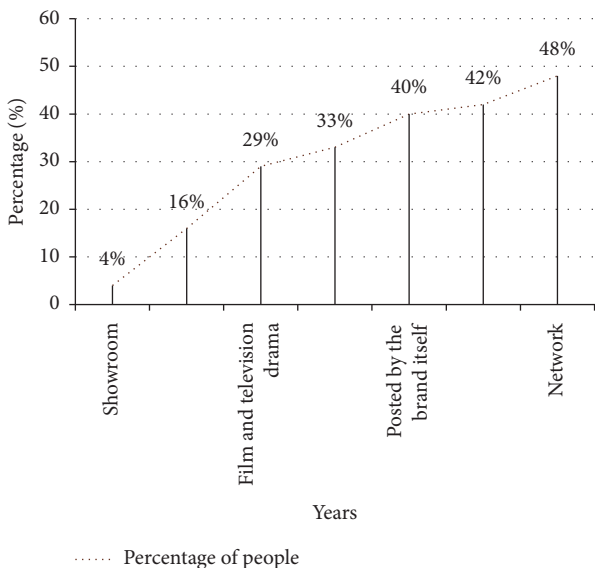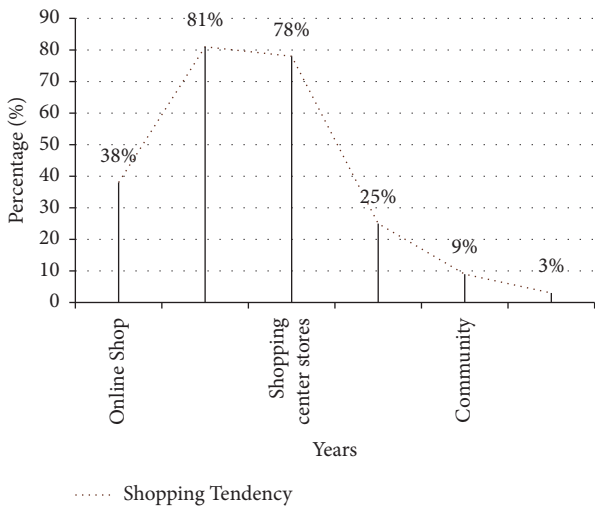
FIGURE 9: Channels for purchasing goods and channels for obtaining fashion information.
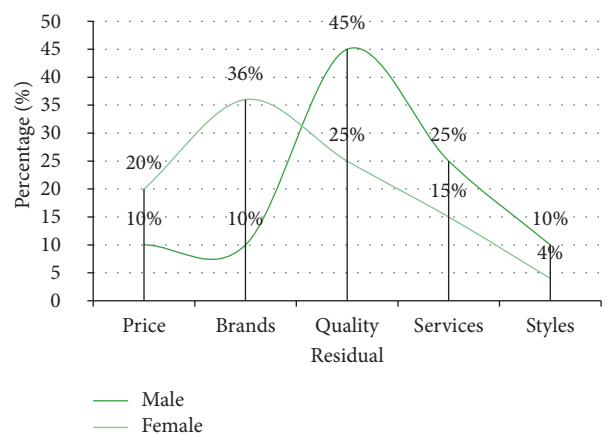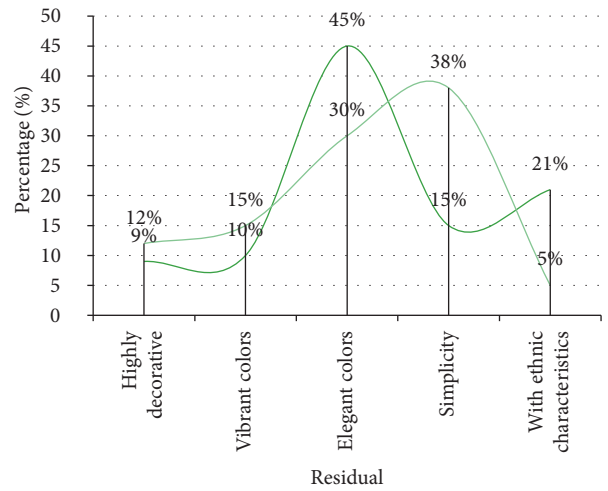
FIGURE 10: Top aspects consumers pay attention to when purchasing a product.

and major brands. Other avenues include friend recommendations, movies, outdoor advertising, and fashion shows.

For the brand business marketing style, there is a large gender difference in the style that the respondents prefer, as shown in Figure 10. The majority of male respondents prefer commercials with simple and elegant colors, accounting for 45%; the majority of female respondents prefer simple-style commercials, accounting for 38%. Regarding the aspects that consumers pay attention to when purchasing products, male customers pay the most attention to the quality of products, and female customers pay the most attention to brands, accounting for 45% and 36%, respectively.

In general, with the development of the times, modern consumers are mainly pragmatic. Consumers have little influence on the company's culture, service concept, store design and layout, and the influence of people around them, as well as advertisements and celebrity endorsements.

5. Conclusions

In today's 21st century, with the continuous development of technologies and applications such as computer networks, mobile Internet, and social networks, computer technology is gradually integrated into people's daily lives, and "data" is also constantly infiltrating people's lives. The promotion of brand value is of great significance to enterprises. On this basis, a decision support system for brand design based on a data warehouse was proposed. Data mining technology was used to obtain design element information, decision support for brand design was provided through human-computer interaction tools, and information sharing, communication, and application were realized. The system not only provided an objective and scientific thinking method and theory for brand design but also filled the gap of research theory using information technology theory to assist brand design work. It organically integrated various information technologies and brand design theories to form an efficient, scientific, and accurate brand design method, and refined design information from the brand design information data from multichannel sources. In the era of big data, a brand design must conduct in-depth research on the social environment, transform design thinking, and use information technology to adapt to current consumption and trends.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. S. Lee and I. Soon, "Taking a bite out of Apple: jailbreaking and the confluence of brand loyalty, consumer resistance and the co-creation of value," *The Journal of Product and Brand Management*, vol. 26, no. 4, pp. 351–364, 2017.

[2] T. Mandler, "Beyond reach: an extended model of global brand effects," *International Marketing Review*, vol. 36, no. 5, pp. 647–674, 2019.

[3] R. Huaman-Ramirez, N. Albert, and D. Merunka, "Are global brands trustworthy?The role of brand affect, brand innovativeness, and consumer ethnocentrism," *European Business Review*, vol. 31, no. 6, pp. 926–946, 2019.

[4] S. Sarkar and S. Banerjee, "Brand co-creation through triadic stakeholder participation," *European Business Review*, vol. 31, no. 5, pp. 585–609, 2019.

[5] G. C. Nobre and E. Tavares, "Scientific literature analysis on big data and internet of things applications on circular economy: a bibliometric study," *Scientometrics*, vol. 111, no. 1, pp. 463–492, 2017.

[6] S. F. Ochoa, G. Fortino, and G. Di Fatta, "Cyber-physical systems, internet of things and big data," *Future Generation Computer Systems*, vol. 75, pp. 82–84, 2017.

[7] W. Jianhua, G. Zhang, and S. Zhengxiang, "Internet of things, big data, intelligent electrical apparatus-the future development of power equipment," *Gaoya Dianqi/High Voltage Apparatus*, vol. 54, no. 7, pp. 0001–0009, 2018.

[8] K. Sugimoto and S. Nagasawa, "Luxury branding: the case of Chanel," *International Journal of Quality and Service Sciences*, vol. 9, no. 3/4, pp. 292–301, 2017.

[9] S. M. Hegner, M. Fetscherin, and M. van Delzen, "Determinants and outcomes of brand hate," *The Journal of Product and Brand Management*, vol. 26, no. 1, pp. 13–25, 2017.

[10] P. Foroudi, K. Dinnie, P. J. Kitchen, T. C. Melewar, and M. M. Foroudi, "IMC antecedents and the consequences of planned brand identity in higher education," *European Journal of Marketing*, vol. 51, no. 3, pp. 528–550, 2017.

[11] M. Deb and A. Agrawal, "Factors impacting the adoption of m-banking: understanding brand India's potential for financial inclusion," *Journal of Asia Business Studies*, vol. 11, no. 1, pp. 22–40, 2017.

[12] W. I. Anthony, H. Fong, and A. Leong, "Predicting event tourists' gambling decision:the cross-level effect of brand equity," *International Journal of Contemporary Hospitality Management*, vol. 31, no. 7, pp. 2951–2969, 2019.

[13] D. Vashist, "Effect of product involvement and brand prominence on advergamers' brand recall and brand attitude in an emerging market context," *Asia Pacific Journal of Marketing & Logistics*, vol. 30, no. 1, pp. 43–61, 2018.

[14] G. Zanete and K. P. Siaw, "The effects of social media brand personality on brand loyalty in the Latvian banking industry: The mediating role of brand equity," *International Journal of Bank Marketing*, vol. 37, no. 6, pp. 1480–1503, 2019.

[15] A. Eldegwy, T. H. Elsharnouby, and W. Kortam, "How sociable is your university brand?An empirical investigation of university social augmenters' brand equity[J]," *International Journal of Educational Management*, vol. 32, no. 5, pp. 912–930, 2018.

[16] L. Hatzithomas, P. Gkorezis, A. Y. Zotou, and G. Tsourvakas, "The impact of atmospherics on WOM about short life-cycle products: the case of motion pictures," *The Journal of Product and Brand Management*, vol. 27, no. 5, pp. 471–483, 2018.

[17] L. Saebom and S. M. Park, "An analysis on sony's failure from the view point of design management," *Journal of Integrated Design Research*, vol. 16, no. 2, pp. 81–90, 2017.

[18] N. Verdonk, J. Wilkinson, J. Culbert, R. Ristic, K. Pearce, and K. Wilkinson, "Toward a model of sparkling wine purchasing preferences," *International Journal of Wine Business Research*, vol. 29, no. 1, pp. 58–73, 2017.

[19] M. Kienzler, "Does managerial personality influence pricing practices under uncertainty," *The Journal of Product and Brand Management*, vol. 26, no. 7, pp. 771–784, 2017.

[20] E. Lee, "Research on the condition of small household's housing and small-housing market for the private Constructors' Targeting strategy," *Journal of the Korean Institute of Interior Design*, vol. 26, no. 5, pp. 117–124, 2017.

[21] J. Tan, D. Xia, S. Dong, H. Zhu, and B. Xu, "Research on pre-training method and generalization ability of big data recognition model of the internet of things," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 5, pp. 1–15, 2021.

Research Article

Machine-Learning-Based Human Fall Detection Using Contact- and Noncontact-Based Sensors

Ayush Chandak [iD],[1] **Nitin Chaturvedi** [iD],[1] **and Dhiraj** [iD][2]

[1]*BITS Pilani, Pilani Campus, Pilani, India*
[2]*CSIR-CEERI, Pilani, Rajasthan, India*

Correspondence should be addressed to Ayush Chandak; ayushchandak876@gmail.com

Automated human fall detection is an essential area of research due to its health implications in day-to-day life. Detecting and timely reporting of human falls may lead to saving human life. In this paper, fall detection has been targeted using machine-learning-based approaches from two perspectives regarding data sources, that is, contact-based and noncontact-based sensors. In both of these cases, various methods based on deep learning and machine learning techniques have been attempted, and their performances were compared. The approaches analyze data in fixed time windows and extract features in the time domain or spatial domain which obtain relative information between consecutive data samples. After experimentation, it was found that the proposed noncontact-based sensor techniques outperformed the contact-based sensor techniques by a margin of 1.82%. After this, it was also found that the noncontact-based sensor techniques outperformed the state of the art of noncontact-based sensor results by a margin of 3.15%. To better suit these techniques for real-world applications, embedded board implementation and privacy preservation of subject by using advanced methods such as compressive sensing and feature encoding need to be attempted.

1. Introduction

The falls are widespread among elderly individuals due to the weakening of body parts which occurs due to aging [1]. Sudden falls can affect not only the elderly but also people of all age groups. Falls can cause injuries ranging from fractures to concussions and, in extreme cases, even death. Due to this, falls have recently been an extensively researched topic, and various automation methods are being studied to detect and analyze them. Thus, there is a growing need to use the latest techniques for the automated detection of falls.

Falls can occur in various different ways. Oneill et al., in their study, classified fall into forward, sideward, and backward [2]. Out of this, it was observed that forward fall was the most prominent fall, with 56.33% of all the falls, with 17.60% of falls being backward and 23.23% of falls being sideward.

Also, the chances of falls increase with increasing age varying from about 26% of people suffering falls at least once

a year in the age group of 65 to 74 to 36.5% in the age group above 85 [3]. Thus, detecting falls among the elderly is more critical due to increased chances of falls and slow recovery rates. The reasons for falls among the elderly can be many, and the consequences can be severe if not detected in time. Unconsciousness resulting from the fall can delay the treatment process; thus, an automated fall detection system can help the concerned people know about the fall quickly. For this purpose, falls can be detected through either contact-based or noncontact-based sensors. Both methods are equally important in different ways.

There are many methods through which fall detection can be approached depending on which sensors are used, like accelerometers, gyroscopes, cameras, and others. In this paper, we have discussed some ways to detect falls using vision-based tools and contact-based sensors. Cameras are becoming very prominent in everyday environments and can be found in various environments such as railways and airports. At the same time, accelerometers are the devices

used for capturing acceleration. When fitted to a body part, they provide information on the acceleration of the body parts. Using the information from data captured on acceleration and the change in acceleration, detection algorithms can be designed and trained to perform automatic fall detection. Thus, these two cases provide very different approaches to fall detection.

Neural networks are becoming very popular in the areas of fall detection due to increased efficiency in the detection process. Convolution neural network (CNN) is the method that specializes in analyzing sequential and visual data [4]. CNNs have been used for tasks such as image classification, natural language processing, and human activity classification. Thus, in this paper, many proposed approaches are designed based on neural networks architectures.

This work has contributed towards (1) detecting falls based on cameras and (2) detecting falls based on smartphone sensors. Various approaches are applied and tested in both cases. Then comparisons of these approaches are presented. Some techniques are also introduced to help reduce computation time and efficiency.

The paper first provides related works in the field and the description of the detection dataset used (UP-Fall). Then, the approaches used for contact-based and noncontact-based fall detection are discussed, followed by the results and analysis. Then, finally, conclusions are drawn.

2. Related Works

Recently, in the field of human activity recognition, many research studies are being done as this is an essential aspect and has a wide variety of use-cases such as healthcare, sports, and elderly care. Of the research works done, many utilized a variety of ways to detect falls, like using classical ML models, which include Support Vector Machine (SVM), K-nearest neighbors, and Random Forest, and neural networks based model. Recently deep learning models have become very popular due to their promising results.

The falls systems were mainly based on contact-based and vision-based sensors. Thakur [5], in his research, analyzed various recent fall detection systems using contact-based, ambiance-based, and vision-based sensors. Many datasets are also increasingly becoming available in both aspects. The two following domains were analyzed for the work being done in automated human fall detection.

2.1. Machine-Learning-Based Fall Detection Using Contact Sensors. This type of sensor relies on sensors being fitted on the subject to detect any sudden motion indicating a fall. They include using sensors such as accelerometers and gyroscopes. These sensors are generally fitted on the watch or mobile phone. These sensors record acceleration of the body parts on which they are fitted. The change in acceleration due to change in the activity changes acceleration is utilized for fall detection.

Many different approaches in this domain have been applied and researched utilizing different aspects. Khojasteh et al. [6], in their research, proposed a fall detection system based on wearable devices like smart wristbands where they utilized a threshold-based solution for this purpose in which if any change in acceleration crossed a certain threshold, then fall signal was generated. Kwolek [7] utilized a similar method based on threshold-based detection on accelerometers. He proposed an innovative solution based on a fuzzy-based system to authenticate the fall event. Şengül et al. [8] used a quaternion algorithm to detect falls and send a signal to the concerned people with the patient location. More recently, fall detection using convolutional neural networks has become popular due to these networks' ability to learn features without any data processing and their robustness and flexibility. Santos et al. [9], in their research, utilized CNN-based models on the accelerometers-based dataset.

2.2. Machine-Learning-Based Fall Detection Using Noncontact Vision Sensors. The major advantage of this type of fall detection is that the user is not required to wear any type of sensor. However, the user is required to be always inside the region of operation of the camera. Also, detecting falls on the border is difficult in these cases [10]. They include the detection of falls using video camera recordings, I.R. sensors, bioradars, and others.

Many different approaches have been attempted in this case utilizing machine learning and deep learning techniques, as well as using different preprocessing techniques like optical flow images, image lightning, and others. Rougier et al. [11], in their study for detecting fall events using a video surveillance camera, used the shape deformation of humans in the video frames during the fall for the detection. Furthermore, shape analysis models were used for training. Capturing human deformation during the fall is a good technique as it is very much noticeable in the fall video frames. Anishchenko et al. [12], in their research, showed the advantages of using bioradar's systems to detect human falls. The research used wavelet transform and neural networks for fall detection. At the same time, Miao et al. [13] utilized CNN-based models to detect human-based activity and further modified them for detecting falls. In a similar approach, the CNN model using 3D ConvNets was approached in this case as well. More recently, latent feature polling was used for detecting falls effectively in ADLs (activities of daily living) [14].

3. Dataset Description

For this research, we have chosen a publicly available dataset, the UP-Fall detection dataset [15], after analyzing the various fall detection datasets [16–20]. Due to variability in subjects, activities, and different lighting conditions, this dataset provided real-life conditions for training and testing. Of all the datasets, the UP-Fall dataset provided more variability in terms of subjects, activities, frontal and lateral views, a greater number of trials, and so forth, and it contained a larger number of videos enough for effectively training the network. Besides that, it also provided data from contact-based sensors like accelerometers, gyroscopes, and

others, which can also be used for trying other approaches to fall detection. The dataset and features are publicly available [21]. The dataset consists of 11 activities performed by 17 young, healthy individuals (8 females and 9 males) in the age group of 18–24 years. Out of the 11 activities performed, 6 consisted of daily human activities, which are lying, jumping, picking up an object, sitting, standing, and walking, and 5 consisted of fall activities, out of which two consisted of falling forward using a hand and knees, two consisted of falling sideways and backward, and one consisted of falling sitting on a chair. Figure 1 describes the positions of some of the contact-based and noncontact-based sensors that were used in the UP-Fall dataset.

For the purpose of the experiment, we have used data recorded at 18 FPS from the two cameras and Inertial Measurement Unit (IMU) sensors fitted on the wrist and right pocket of the subject. We aim to implement a system to distinguish between fall and nonfall activities using data from contact-based sensors such as accelerometers as well as noncontact-based sensors such as cameras.

4. Methodology and Implementation

Human fall detection can be captured through a variety of methods. The analyzed methods were chosen such that they are not much affected by the environmental factors and can be applied to any general setting (not requiring any special arrangements). The following section describes the methods and the architectures that were implemented for contact-based and noncontact-based fall detection (Figure 2).

4.1. Detection Using Contact-Based Sensors. Data from the triaxial accelerometer was utilized in this case for fall and nonfall prediction. Two 3-axial accelerometer sensors were fitted on the subjects in the right pocket and on the right wrist simulating the wristwatch and cell phone, respectively. Since the accelerometers are kept very close to the subject, predicting the fall using them is very institutive for automated detection. Machine learning architectures and a 1D CNN-based approach were deployed for this purpose.

Firstly, the 3-axial data received from the accelerometers was trimmed to 3 seconds in order to capture the fall and nonfall parts effectively. To standardize the dataset, the data points were subtracted from the mean and divided by the standard deviation. Afterward, continuous 1.5 seconds of fixed arrays were drawn from the normalized data, which were then appended one after the other in 3rd dimension order. The obtained flow stacks were passed through the classifier (ML models and 1D CNN) to predict the fall event. The overall process outline can be seen in Figure 3. For the experiments, the dataset was divided into 70% for training and 30% for testing.

4.1.1. Machine Learning Methods. In this paper, the following machine learning classification methods were implemented for the detection:

(1) Support Vector Machine (SVM): This method differentiates between the classes by creating a hyperplane and mapping the inputs on it. The hyperplane is further optimized by training and then acts as a decision boundary for prediction. The target of the algorithm is to maximize the gap between the data points and the hyperplane. SVM is a very popular ML algorithm and is used extensively in fall detection systems [22].

(2) Random Forest (RF): This method is a widely used machine learning algorithm that works on the principle of ensemble learning. It works by creating an ensemble of decision trees through which the input is passed, and, on the basis of the mode of the responses received, the output of the algorithm is computed. The greater number of trees leads to better accuracy and reduces risks of overfitting [23].

(3) Multilayer Perceptron (MLP): This method is a feedforward method passing inputs through the multiple layers before computing the final output. Each node has a neuron that utilizes a nonlinear activation function. It uses the back-propagation technique for the training of the model [24].

The parameters used for the machine learning architecture used are described in Table 1.

4.1.2. 1D CNN Network. CNN is very good for classifying sequential data. Thus the data from the 3-axial accelerometers is a good match for applying CNN-related methods. Lee et al. [25], in their research, analyzed the benefits of using the 1D CNN process for fall detection. In the research, 1D CNN was shown to be better than most of the traditional ML methods when using time-sequential data.

The convolutional and pooling layers used in 1D CNN are one-dimensional with height and width of 1. The algorithm that was used for this purpose is described in Figure 4. It is a simple convolution method with 1st layer being convolutional followed by max pooling and then one more convolutional layer followed by a softmax activation function for calculating probabilities. The network was trained for 500 epochs.

4.2. Detection Using Noncontact-Based Sensors. RGB camera recordings captured from the frontal and the lateral side of the subject were utilized in this method for predicting the fall and nonfall events. Additionally, the RGB images were also converted to optical flow images. Optical flow images are useful in removing the background noises as they only capture the motion between the consecutive frames, as shown in Figure 5. Both RGB and optical flow images were attempted for the detection. Figure 6 describes the stepwise procedure that was used for training the dataset. Overall, it consisted of 5 steps which were dataset collection, segmentation, feature extraction, training of the model, and classification. These steps are explained in the following sections.
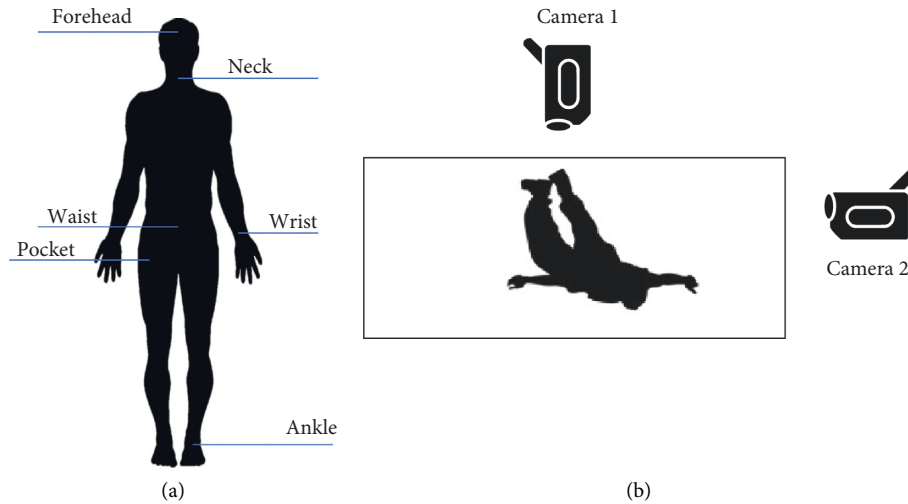
FIGURE 1: Position of the sensors. (a) Contact-based sensors located on the human body. (b) Video camera location layout for the lateral and frontal views.
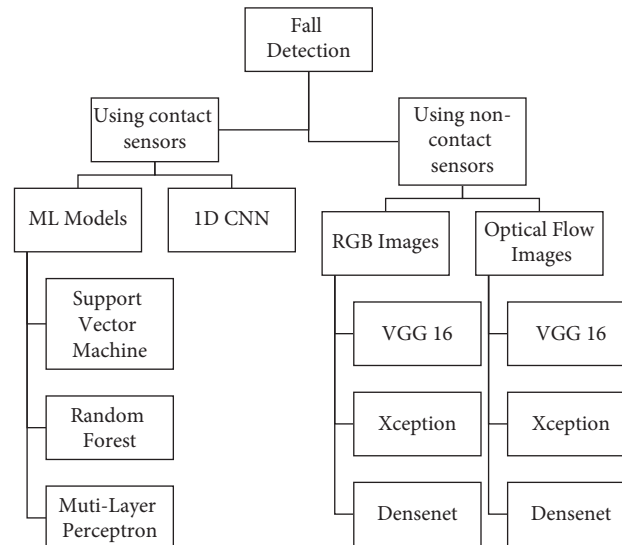
FIGURE 2: Overview of the methods and architecture used in the paper.
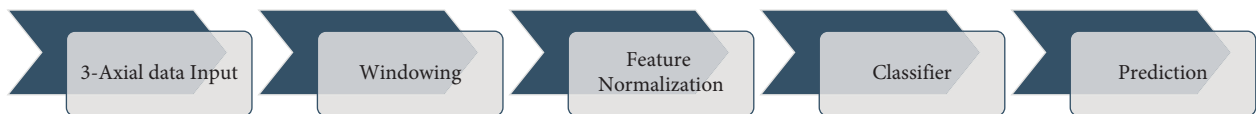
FIGURE 3: The process outline used for fall detection using contact-based sensors: 3-axial data input from accelerometers is processed and then passed through the classifier, which identifies whether there is a fall or not.

4.2.1. Dataset Used. Camera recordings from the UP-Fall detection dataset were used as the data input. Videos were captured from the frontal and lateral sides of the subject at 18 FPS. It comprised 17 subjects performing 11 activities in three trials, as shown in Table 2. Out of the 11 activities, 5 were classified into fall and 6 into nonfall category.

To keep the computational costs low, the resolution of the images was reduced to 224×224 from 640×480. Moreover, the recordings were trimmed to 50 frames each,

given that 50 frames are sufficient for identifying any fall or nonfall event. The dataset was further divided into training set and testing set in an $80:20$ ratio.

4.2.2. Segmentation. As observed from the dataset, detecting any fall event does not require more than 20 frames. Using this observation, the sliding window segmentation technique was attempted to capture video frames containing only the

TABLE 1: Parameters used for ML methods.

Architecture used	Parameters
SVM method	Kernel = "radial basis function"
	$C = 1$
	Gamma = 0.1
	Tolerance = 0.001
RF method	Estimators = 100
	Criterion = "Gini"
	Sample split = 2
	Min. sample leaf = 1
	Max. depth = none
MLP method	Activation = "ReLU"
	Solver = "Adam"
	Max. number of iterations = 200
	Tolerance = 0.0001
	Shuffle = true
	Initial learning rate = 0.001
	Beta1 (exponential rate decay) = 0.9
	Beta2 (exponential rate decay) = 0.999
	Epsilon (numerical stability measure) = $10-8$
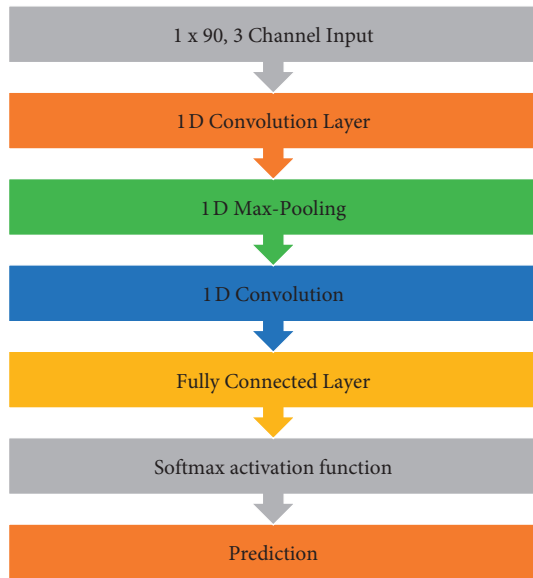
FIGURE 4: Architecture used in case of 1D CNN for sensor-based training of the model.

fall event, if present. So, out of the N frames, if L is the length of the window used, then using a sliding window of 1 frame, $N - L + 1$ windows were captured, as shown in Figure 7.

4.2.3. Feature Extraction.
The target of this step is to get the most important features that are relevant to the experiment. This part involves removing the noise from the data and capturing the part that is most relevant to the training of the architecture. The noise can be any repetitive data or unwanted data points that can affect the convergence of the model.

For vision-based approaches, the optical flow method delivers good information for clear movements between the images and removing any background noises [26]. Thus, for capturing the motion and removing the background noises, the effect on the model by converting the RGB images to optical flow images was analyzed. For this purpose, out of 50 frames, the difference in the intensity of consecutive frames was taken. Thus, starting from the 1st frame, the difference between the 1st and 2nd was taken, followed by that between the 2nd and 3rd, and so on till the 49th and 50th frame. Thus, in this way, we obtained 49 optical images representing the activity. Figure 6 describes the effect of converting an RGB image to an optical flow image in the case of the UP-Fall dataset.

4.2.4. Model and Training.
For training the model with video recordings as the input, 3D convolutional neural networks offer the capability to capture spatial information in time by stacking consecutive video frames. Tran et al. [27], in their study, analyzed the benefits of using 3D ConvNets when using video datasets for training. By using stacks of consecutive frames as the input, they allow capturing the features that take the spatial video frame data and motion between the frames into account. Due to this edge offered in the case of 3D ConvNets, they are increasingly becoming important in the field of human activity recognition. Thus, in this paper, 3D ConvNets were utilized for training the models in the case of noncontact-based sensors.

The training steps used for building the network are as follows:

(I) The model was trained on the ImageNet dataset [28]. ImageNet is an extensive image dataset that comprises more than 14 million images and 1000 classes such as animals, sports, and daily human activities. This helped the model to learn basic generic characteristics about the RGB images, such as identifying border, texture, and color.

(II) From the CNN model trained on the ImageNet dataset, the input layers were modified to take the input as $224 \times 224 \times 20$, where 224×224 is the size of the image and 20 is the stack size. Then the model was fine-tuned on the UCF101 fall dataset [20], which comprises 13,320 videos of 101 different human activities. It helped the model to learn the features of generic human actions. To avoid biasing errors, the fall video frames were reduced in order to contain falls only. 3D ConvNet was used here for training the models.

(III) In the final step, the obtained model was fine-tuned over the UP-Fall dataset.

The following experiments were done to decide the configuration to be used for the training phase.

(1) Learning Rate. It is a very important hyperparameter that decides how the model weights are updated during the training. Deciding on a proper learning rate is very important as very small learning rates can make the training too slow, and higher values can make it too fast to converge the model optimally. It was observed from a series of

RGB Images

Optical Flow Images
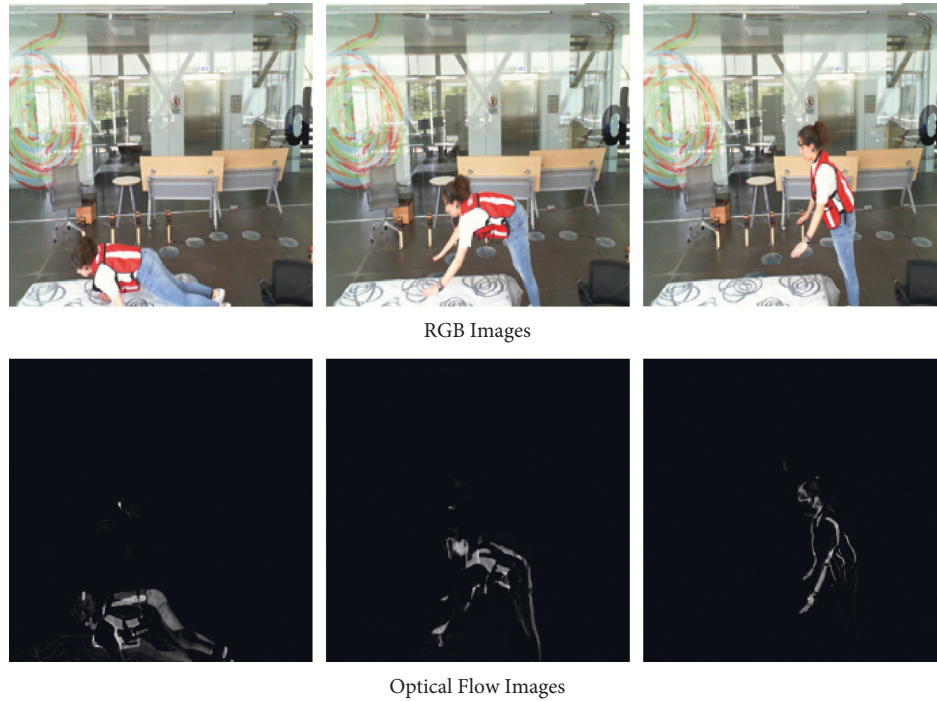
FIGURE 5: RGB and its optical flow image equivalent. Optical flow images help in removing background noises and capture only the motion.
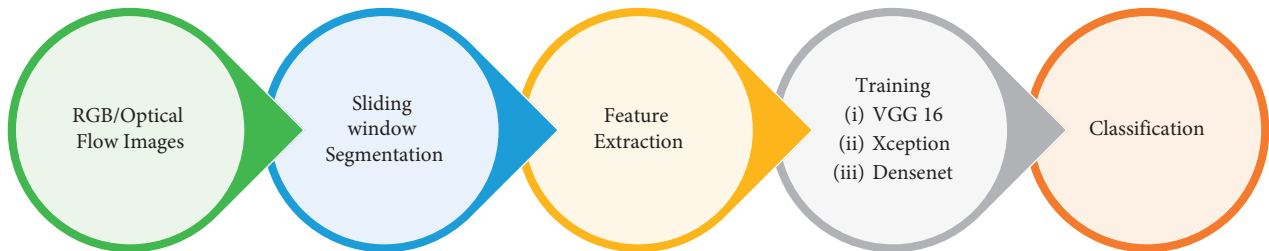
FIGURE 6: Process pipeline used for fall detection using contact-based sensor.

TABLE 2: Duration of activities and their classification [15].

Sr. no.	Description	Duration (in seconds)	Classification
1.	Falling forward with knees	10	Fall
2.	Falling backward	10	Fall
3.	Falling forward using hands	10	Fall
4.	Falling sideward	10	Fall
5.	Falling sitting on a chair	10	Fall
6.	Standing	60	Nonfall
7.	Walking	60	Nonfall
8.	Sitting	60	Nonfall
9.	Jumping	30	Nonfall
10.	Laying	60	Nonfall
11.	Picking	10	Nonfall

experiments that the network provided the best results when the learning rate was between 10^{-3} and 10^{-4}. Choosing higher or lower values affected the accuracy, as lower learning made the system too slow and higher learning made too big steps which made it far away from achieving proper optimization in those cases.

(2) Minibatch Size. Li et al. [29], in their study, showed that using minibatch size properly with proper experimentation yields better results than the standard gradient descent. Minibatch gradient is used to split the dataset into batches, and then the training of the network is done in batches; that is, model weights are updated in the batches. Due to a greater
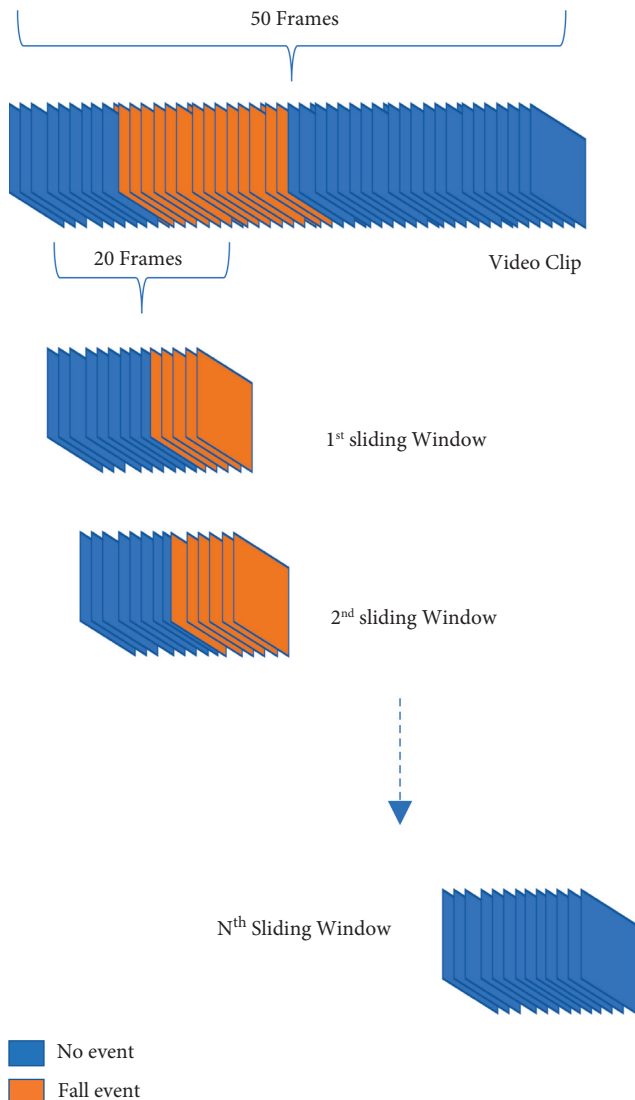
FIGURE 7: Sliding window technique used to capture stacks of consecutive frames. Here frames in orange color are the ones that contain events important for the detection.

number of updates done using minibatch, a good convergence is obtained by using minibatch gradient descent. We tried batch sizes in powers of 2. It was observed through a series of an experiment that a batch size of 1024 was giving the best results.

(3) Number of Epochs. Epoch is the number of times the algorithm passes through the training dataset. It is an important parameter for balancing between the time taken for training (with greater number of epochs, higher computation is required) and the saturation obtained. The number of epochs used was 500 in all the experiments, as proper saturation was seen to be obtained within the 500 epochs.

One more important step in the training process is batch normalization. Ioffe et al. [30], in their research, explained that making the batch normalization method part of the minibatch training process can make the training process much faster. It was further shown that it eliminates the need for using dropout rates to a great extent. The activation function ReLU (Rectified Linear Unit) was used in the experiments with batch normalization to avoid overfitting and faster training.

After fixing the hyperparameters, three different architectures backbones experimented with the 3D ConvNet to fine-tune on UP-Fall dataset. These three were chosen to consider the different types of diverse combinations. In two phases, the training of the model was done. Firstly, RGB images were trained on the architectures in two parts, initially on the frontal view and then on the lateral view. The optical flow images were trained similarly. The architecture used was as follows:

(1) VGG-16 [31]. It is a very popular CNN architecture that is widely used for classification purposes. It consists of 16 layers for training. It uses 3×3 convolutional kernels and 2×2 pooling kernels. The consistent arrangement of these layers is followed throughout the architecture. The input shape of 224×224 was used in the architecture. Figure 8 describes the layers of the VGG-16 architecture used.

(2) Xception [32]. Xception is a CNN architecture that is 71 layers deep. With modified depth-wise separable convolution, it is a modification of Inception v3 with a parameter that is almost equal to Inception v3. Figure 9 gives an overview of the Xception architecture implemented by showing some of the layers presented.

(3) DenseNet [33]. DenseNet is a CNN architecture in which the layers are densely concentrated, which makes its training time somewhat more than other architectures, but, at the same time, it leads to better convergence of the model. There are many different versions of DenseNet, which vary by the difference in the layer depth. In the experimentation, DenseNet201 with 201 layers was used. The major advantage of using DenseNet as compared to other architectures is that it is more efficient than the other models in capturing low-level features of the images. The overview of the DenseNet architecture can be seen in Figure 10.

5. Results and Analysis

For evaluating the performance of models, developed accuracy, sensitivity, specificity, precision, and $F1$-Score were used. Accuracy defines the correctness of the classification model. Sensitivity and specificity define the ability of the system to detect the falls and nonfall events, respectively. Precision predicts the rate of happening of positive events. $F1$-Score is another measure of accuracy focusing on precision and sensitivity.

Equations (1)–(5) define the formulas used to calculate them. Here, F.P. denotes False Positive, F.N. denotes False Negative, T.P. denotes True Positive, and T.N. denotes True Negative. These values are defined as follows:
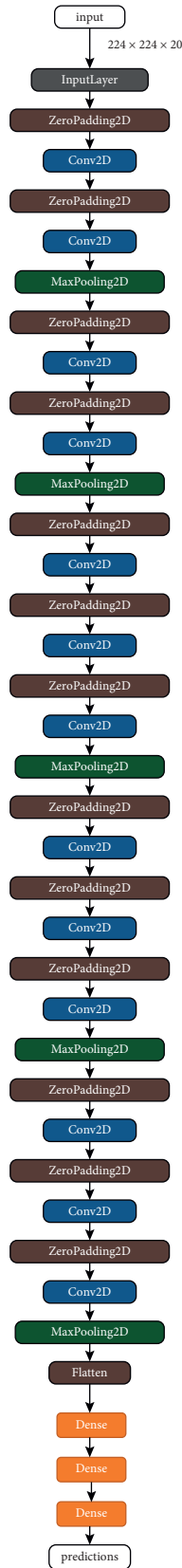
FIGURE 8: VGG-16 architecture: padding layers in brown, convolution layer in blue, pooling layers in green, and dense layers in orange.
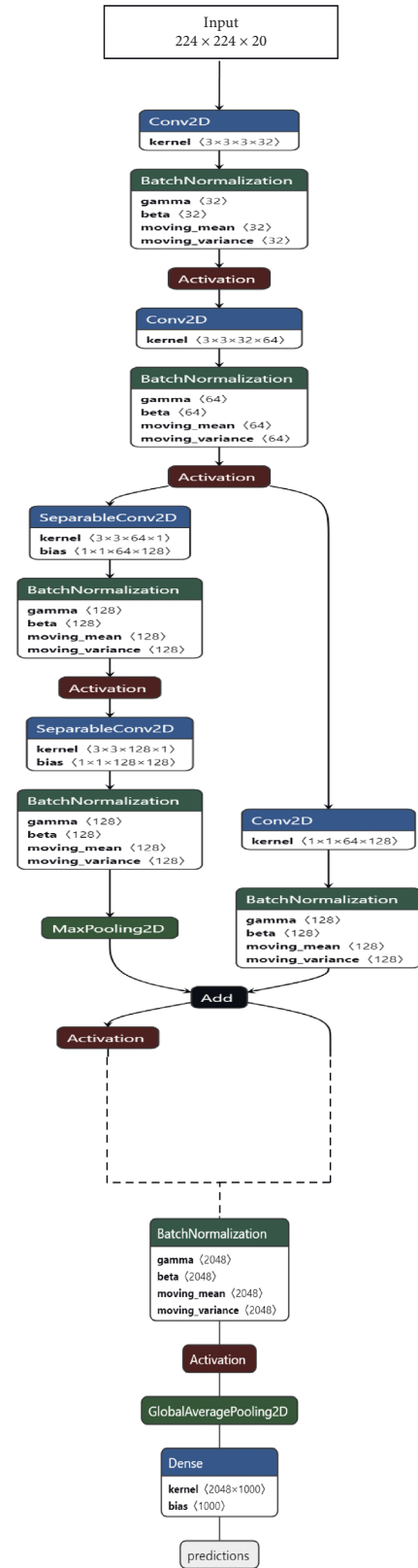
FIGURE 9: Overview of the Xception architecture implemented. Due to large number of layers present, some of the layers from the start and end are shown here.
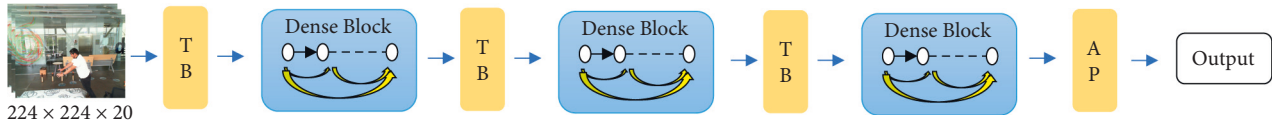
FIGURE 10: Overview of the DenseNet architecture. T.B. is transition block, and A.P. is average pooling and connected block.

$$\text{Accuracy} = \frac{(T.P + T.N)}{T.P. + T.N. + F.P. + F.N}, \tag{1}$$

$$\text{Sensitivity} = \frac{T.P.}{(T.P. + F.N.)}, \tag{2}$$

$$\text{Specificity} = \frac{T.N.}{(T.N. + F.P)}, \tag{3}$$

$$\text{Precision} = \frac{T.P.}{(T.P. + F.P)}, \tag{4}$$

$$F1 - \text{Score} = 2.\frac{(\text{Precision.Sensitivity})}{(\text{Precision} + \text{Sensitivity})}. \tag{5}$$

The models trained on similar datasets were compared based on these parameters. All the experiments related to CNN architectures were performed on Python v3.7.3 with Keras v2.2 [34] framework using TensorFlow as backend, and ML methods were trained using sklearn library [35]. The hardware used for the experimentation includes 2 Nvidia GTX 1080 Ti graphic cards with 32 GB physical RAM and Intel i7 3.6 GHz processor.

5.1. Using Contact-Based Sensors

5.1.1. Using ML Methods. The 3-axial data obtained from the accelerometer fitted on the right wrist of the subject was used for training the model on 3 machine learning methods, which were Support Vector Machine, Random Forest, and Multilayer Perceptron. The data obtained from accelerometers captured acceleration along the x-axis, y-axis, and z-axis in continuous time frames. The data was further divided into training set and test set in the ratio of $70 : 30$.

The results obtained from the three algorithms are summarized in Table 3. It can be seen that the RF method performed the best among all the machine learning methods on all the performance metrics. The results also suggest that ML methods can provide significant efficiency on the sequential data stream provided by the accelerometers. To improve on the performance, we further tried the 1D CNN method.

5.1.2. Using 1D CNN Method. Convolutional neural network methods are very good at training sequential streams of data; here 1D CNN-based method was implemented. In 1D CNN, data streams are passed after converting them into 1D arrays and passing them through multiple 1-dimensional layers for training. The training was performed for 500 epochs, with a decent gradient optimizer with a learning rate of 10^{-4}.

Here the descent gradient optimizer was used with a negative log-likelihood cost function, as shown in the following equation:

$$\text{Loss}(x) = -\log(x). \tag{6}$$

Figure 11 shows the accuracy/loss plots obtained in this case. An accuracy of 98.07% was obtained, shown in Table 3. Thus, it was observed to be better than the other implemented ML algorithms.

5.2. Using Noncontact-Based Sensors. The model's training was done in two phases through RGB images and optical flow images. RGB images included the frames obtained directly from the camera, while optical flow images were obtained to capture the motion between the consecutive frames. Then, in both cases, the dataset was divided into training set, 80%, and test set, 20%.

The K cross-validation approach was used to remove the biasing effect, where K refers to the number of different types of groups the data was split into. It is a beneficial technique as it helps get the models variety of combination training and testing datasets. The value of K was chosen to be 5 as selecting a higher value of K leads to overfitting, and a lower value of K is not high enough to remove the biasing effect, so the value of 5 seemed optimal.

5.2.1. Using RGB Cameras. In this experiment, three architectures were trained, namely, VGG-16, Xception, and DenseNet, on lateral and frontal views of the subject. The learning rate was set at 10^{-6}, with batch normalization with ReLU as an activation function. The network was then trained for 500 epochs.

Tables 4 and 5 show the summary of the performances of all the three architectures in lateral and front views, respectively. Figures 12 and 13 highlight the accuracy/loss plots for lateral and frontal views, respectively.

The frontal view was seen to be slightly better than the lateral view; this is because the subject occupies more part of the image in the former case. It can be observed that DenseNet outperformed other architectures with an accuracy of 98.41% in the case of lateral camera recordings and with an accuracy of 99.85% in the case of frontal camera recordings. So, it can be concluded that DenseNet has better performance than the other architectures in the case of fall detection.

When compared with the other vision-based fall detection methods reported in the literature [14, 36], which also implemented CNN-type architectures on UP-Fall detection dataset, as shown in Table 6, it can be seen that the proposed methodology produced results that are comparable with the state-of-the-art methods.

TABLE 3: Summary of results of ML methods implemented.

Architecture used	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
SVM method	94.00	93.54	93.00	94.00	93.26
RF method	96.875	96.77	96.77	96.96	96.77
MLP method	93.75	93.54	93.54	93.93	93.53
1D CNN	98.30	1	97.14	1	98.55
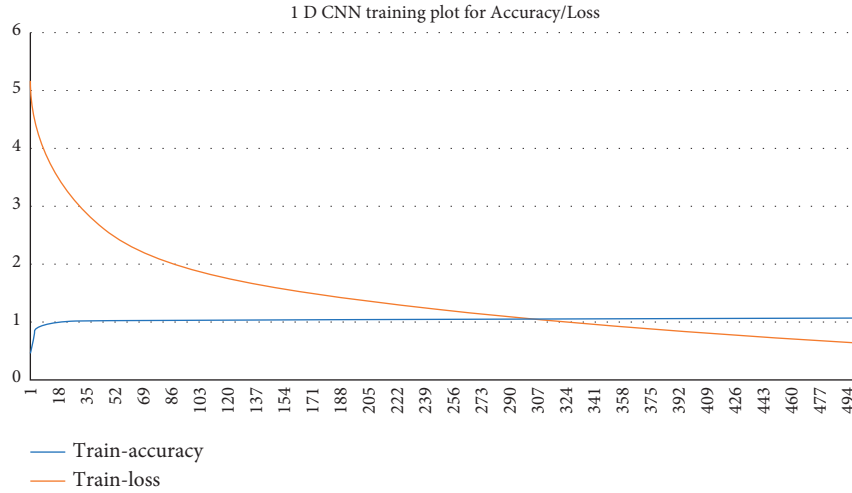
FIGURE 11: Training accuracy/loss plot obtained in the case of 1D CNN.

TABLE 4: Summary of results of architectures implemented in case of RGB images from a lateral view.

Architecture used	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
VGG-16	97.19	98.25	96.09	98.29	97.16
Xception	98.04	96.90	99.26	96.82	98.07
DenseNet	98.41	98.53	98.29	98.53	98.41

TABLE 5: Summary of results of architectures implemented in case of RGB Images from a frontal view.

Architecture used	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
VGG-16	96.58	94.26	98.78	94.39	96.65
Xception	97.80	96.44	96.34	94.39	97.02
DenseNet	99.85	99.87	99.82	99.88	99.84

(a)

(b)

FIGURE 12: Continued.

(c)
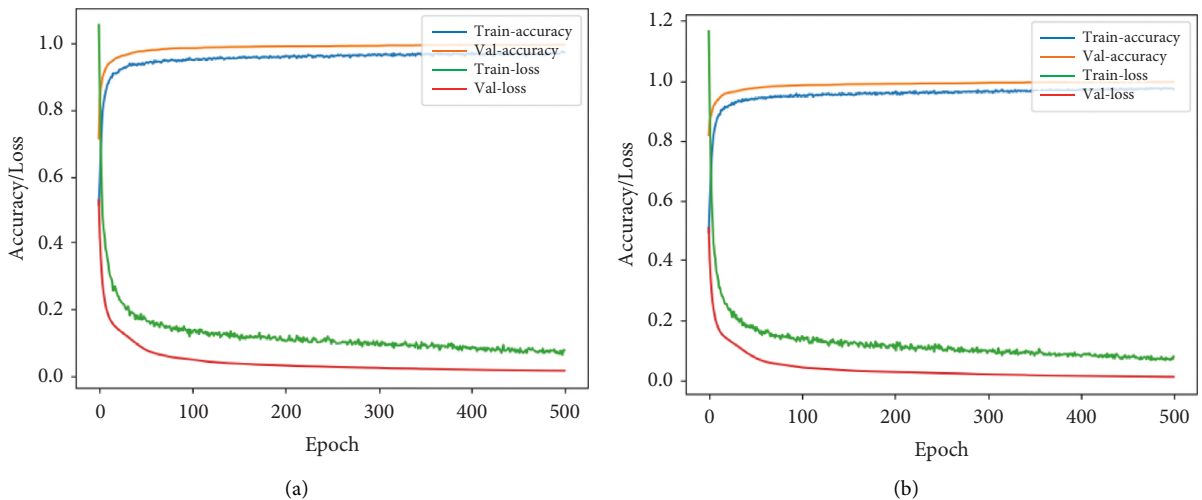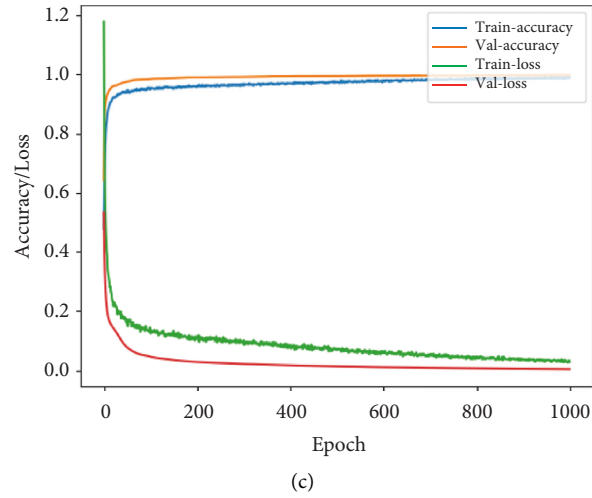
Figure 12: Accuracy/loss training plot for the case of RGB camera recordings in case of lateral view for the three architectures. (a) VGG-16. (b) DenseNet. (c) Xception.

(a)

(b)

(c)
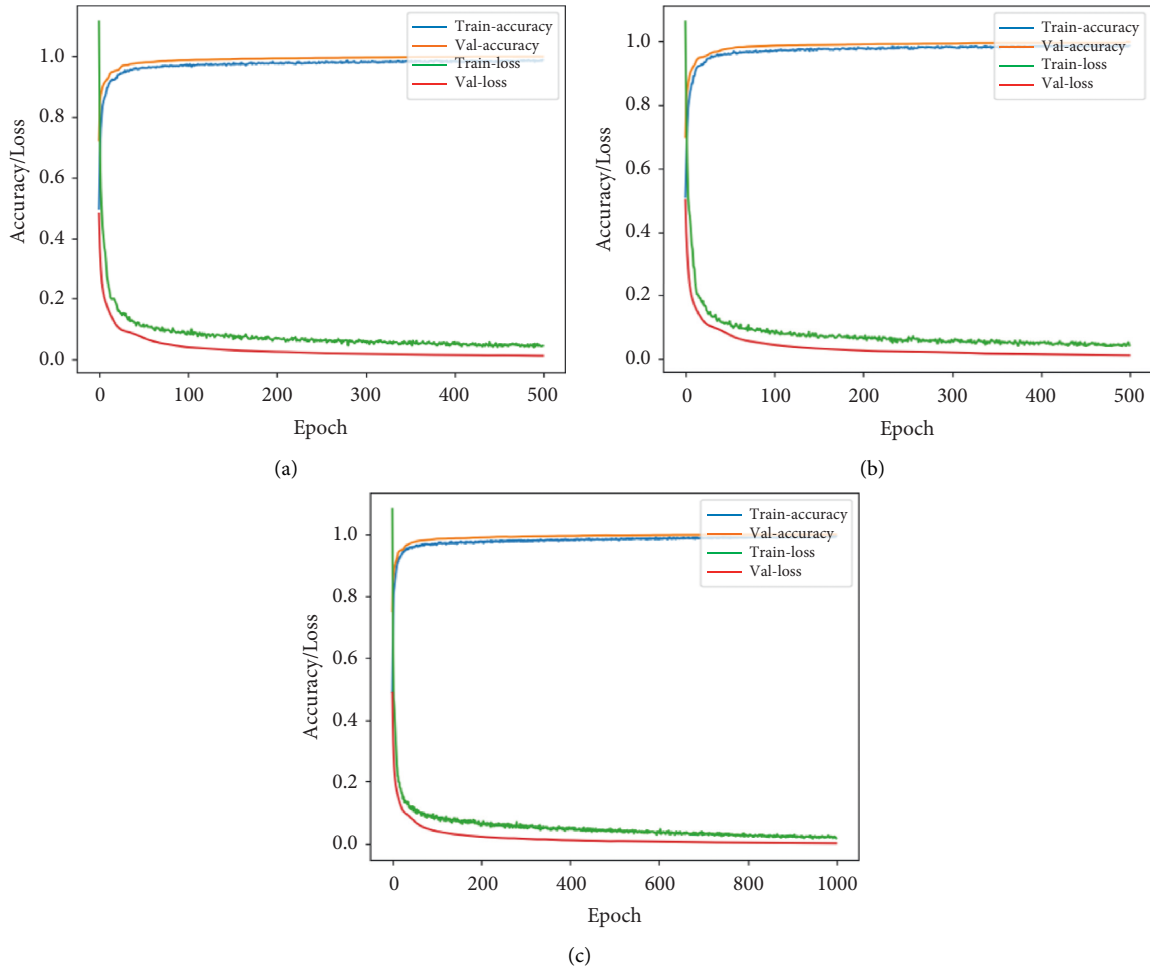
Figure 13: Accuracy/loss training plot for the case of RGB camera recordings in case of frontal view for the three architectures. (a) VGG-16. (b) DenseNet. (c) Xception.

TABLE 6: Comparison with the literature for fall detection from the frontal view.

Proposal	Dataset	Method	Accuracy (%)
Espinosa et al. [36]	UP-Fall	CNN	95.25
Yadav et al. [14]	UP-Fall	CNN	96.70
Proposed method (DenseNet)	UP-Fall	CNN	99.85

TABLE 7: Summary of results of architectures implemented in case of optical flow images from the lateral view.

Architecture used	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
VGG-16	94.125	90.57	98.5	89.75	94.37
Xception	96.125	93.617	99.00	93.25	96.23
DenseNet	96.00	94.66	97.5	94.5	96.05

TABLE 8: Summary of results of architectures implemented in case of optical flow images from the frontal view.

Architecture used	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
VGG-16	92.625	95.46	89.50	95.75	92.38
Xception	93.625	96.28	90.75	96.50	93.43
DenseNet	96.00	96.46	95.50	96.50	95.97

5.2.2. Using Optical Flow Images. Optical flow images are constructed based on the information between the consecutive frames. They capture the motion, if any, between the continuous frames and remove background noises. Tables 7 and 8 show the performances of the architectures in the case of optical flow images of lateral and frontal side views, respectively. They compare the architectures in both the lateral and frontal views based on accuracy, precision, sensitivity, specificity, and F1-Score.

The learning rate was kept at 10^{-6}, batch normalization with ReLU as activation function was used, and again it was observed and concluded that the DenseNet architecture gave better performance than the other architectures in both the lateral and frontal views. Also, it was observed that the frontal view was giving better performance when compared to the lateral view due to the subject occupying more space in the former case.

In Figures 14 and 15, training accuracy/loss plots are shown for all three architectures in the case of lateral and frontal views, respectively.

Overall, two approaches based on the contact- and noncontact-based sensors data were attempted to detect falls. Analysis of the results suggests that the 1D CNN method, with an accuracy of 98.07%, outperformed the traditional machine learning methods in the case of contact-based sensors.

It can also be seen that dataset preprocessing steps, which include segmentation and feature extraction, are very important steps for improving the performance of the models by leading to better convergence. Similarly, with an accuracy as high as 99.85% in the case of noncontact sensors, DenseNet outperformed the other architectures and it is shown that very good predictions can be made using camera recordings as well.

The two methods using contact- and noncontact-based sensors discussed above were totally independent. The key advantages of the approaches are discussed here. First, user privacy is a very important concern when going for fall detection based on video recordings. To tackle this, the previous recording can be discarded immediately from the system after processing. In sensor-based detection, only data from the accelerometer is required, so there are not many issues of personal privacy there. For further improving privacy through video camera recording, an optical flow camera-based approach was also discussed, which allows only relevant information. Thus, in both approaches, privacy of an individual can be preserved. Second, in sensor-based fall detection using 1D CNN, the proposed architecture is very efficient in terms of high efficiency at a very low computation cost due to the simplicity of the architecture.

The architectures were implemented on the UP-Fall detection dataset, which includes 17 individuals performing 11 activities. Out of 11 activities, 5 activities were classified as fall, and the other 6 were classified as nonfall. The actors in the dataset were aged 18–24 years. Since the fall activity is only composed of motion, the approaches can be easily transferred to people of all age groups. So, we think these approaches can be extended to apply to real-life conditions as well.

It is also very important to discuss the limitations of the proposed approaches. There are many activities that can be classified into falls. Detecting different falls happening for a very short period can be difficult. For achieving very good efficiency in the case of camera-based detection, the computation cost of the system required will be very high. For real-time fall detection, the system processing speed would be required to be on par with the model processing time as there should not be lag; otherwise, that lag would keep on accumulating to create a huge difference in time and memory requirements.

One important difference between the two approaches is that fall using noncontact-based sensors has low computation cost, and the user has the flexibility of movement. Meanwhile, through camera-based fall detection, the user is required to be inside the frame of the recording. Even falls happening on the edges could be difficult to recognize.

(a)

(b)

(c)
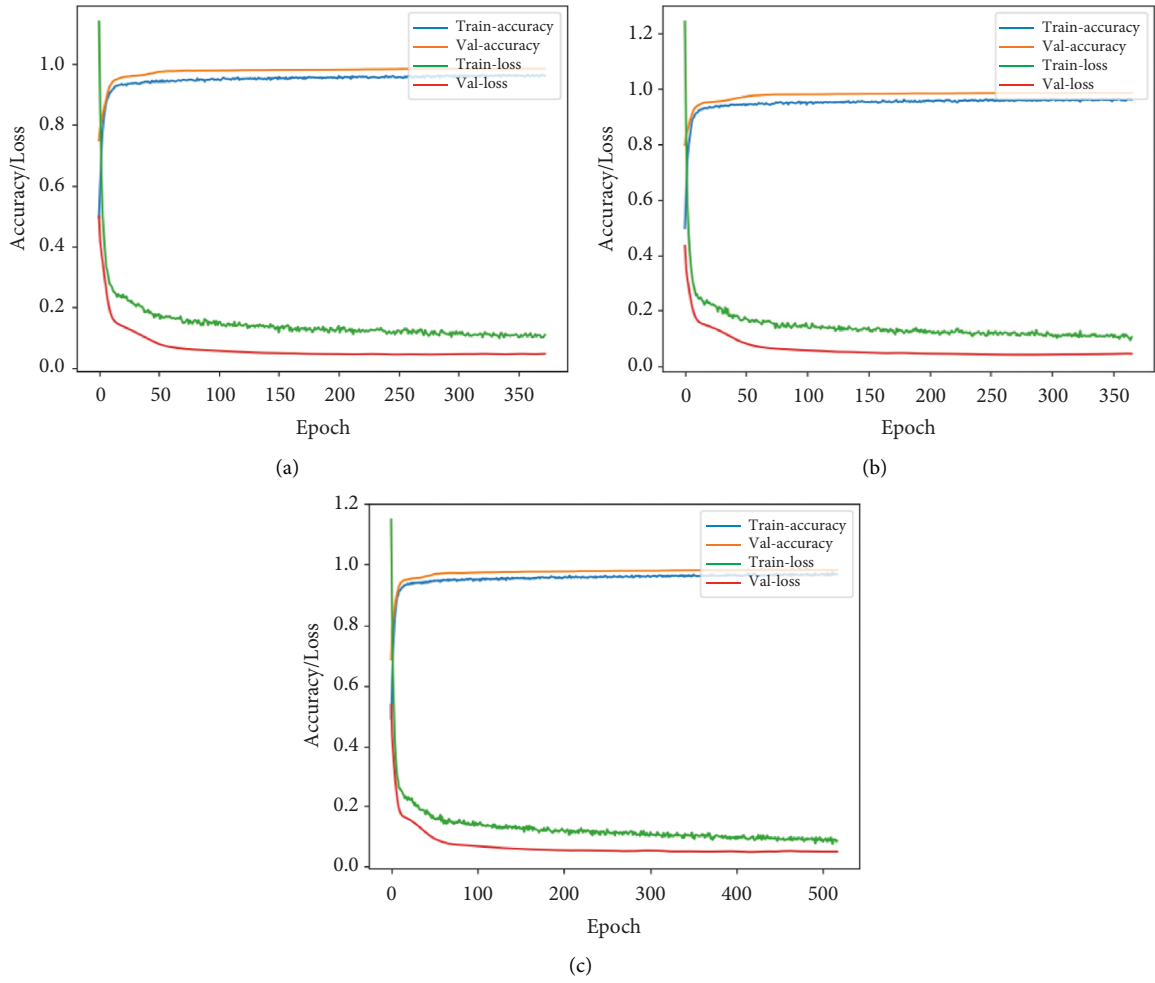
FIGURE 14: Accuracy/loss training plot for the case of optical flow images in case of lateral view for the three architectures. (a) VGG-16. (b) DenseNet. (c) Xception.

(a)

(b)
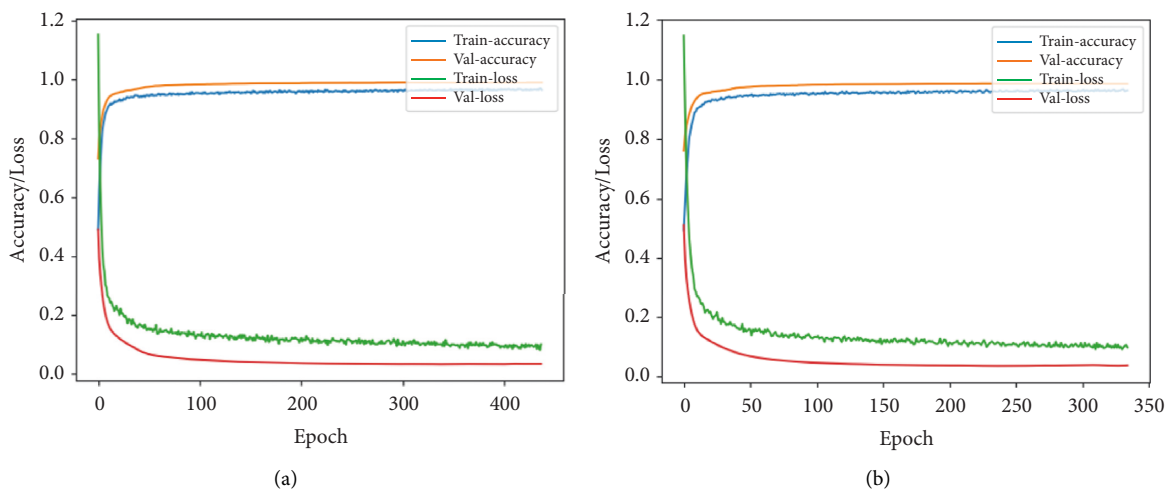
FIGURE 15: Continued.

(c)
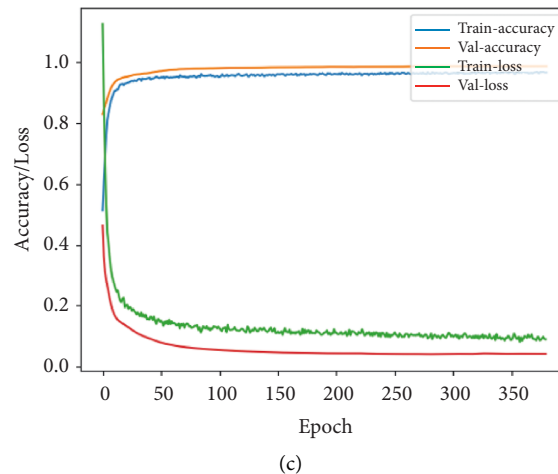
FIGURE 15: Accuracy/loss training plot for the case of optical flow images in case of frontal view for the three architectures. (a) VGG-16. (b) DenseNet. (c) Xception.

6. Conclusions

In this paper, we presented two approaches for fall detection which were using video camera recordings and accelerometer sensors that are commonly found in smartphones. These were done using machine learning and deep learning methods. In the case of contact-based sensors, 1D CNN and machine learning methods were studied. Then, in the case of noncontact-based sensors, 3D ConvNets were implemented utilizing 3 different CNN-based architectures: VGG-16, Xception, and DenseNet.

From the conducted experiments, various things can be concluded. In the case of vision-based sensors, from the three different architectures attempted for implementation, VGG-16 was found to be computationally faster (due to only 16 layers) with good accuracy and comparable with the literature. For further improving on the accuracy, DenseNet is suggested, which performed best among the architectures, but this was obtained at the cost of increasing computation due to a greater number of densely connected layers. In the case of contact-based Sensors, the proposed 1D CNN architectures were shown to outperform ML methods. Both methods have their particular advantages in various conditions. Particularly, sensor-based detection showed various advantages, which include user privacy, no restricted area for movement, and simple architecture.

As can be seen from Table 6, our proposed methodology outperforms the existing architectures with an accuracy of 99.85% as compared to 96.70% of the current state-of-the-art methods. Further improvements in this research should be directed towards embedded board implementation and privacy preservation of the subject by using advanced methods such as compressive sensing and feature encoding.

Data Availability

The UP-Fall detection data used to support the findings of this study have been deposited in the UP-Fall repository (https://sites.google.com/up.edu.mx/har-up/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. F. d. Souza, R. E. A. Batista, C. R. V. Camapanharo, P. C. P. d. Costa, M. C. B. T. Lopes, and M. F. P. Okuno, "Factors associated with risk, perception and knowledge of falls in elderly people," *Revista gaúcha de enfermagem*, vol. 43, Article ID e20200335, 2022.

[2] T. W. O'neill, J. Varlow, A. J. Silman et al., "Age and sex influences on fall characteristics," *Annals of the Rheumatic Diseases*, vol. 53, no. 11, pp. 773–775, 1994.

[3] M. C. Nevitt, S. R. Cummings, and E. S. Hudes, "Risk factors for injurious falls: a prospective study," *Journal of Gerontology*, vol. 46, no. 5, pp. M164–M170, 1991.

[4] S. Albawi, M. Tareq Abed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proceedings of the 2017 international conference on engineering and technology (ICET)*, pp. 1–6, Ieee, Antalya, Turkey, August 2017.

[5] N. Thakur and C. Y. Han, "A study of fall detection in assisted living: identifying and improving the optimal machine learning method," *Journal of Sensor and Actuator Networks*, vol. 10, no. 3, p. 39, 2021.

[6] S. Khojasteh, J. Villar, C. Chira, V. González, and E. de la Cal, "Improving fall detection using an on-wrist wearable accelerometer," *Sensors*, vol. 18, no. 5, p. 1350, 2018.

[7] B. Kwolek and M. Kepski, "Fuzzy inference-based fall detection using kinect and body-worn accelerometer," *Applied Soft Computing*, vol. 40, pp. 305–318, 2016.

[8] G. Şengül, M. Karakaya, S. Misra, O. O. Abayomi-Alli, and R. Damaševičius, "Deep learning based fall detection using smartwatches for healthcare applications," *Biomedical Signal Processing and Control*, vol. 71, Article ID 103242, 2022.

[9] G. L. Santos, P. T. Endo, K. H. de C. Monteiro, E. da S. Rocha, I. Silva, and T. Lynn, "Accelerometer-based human fall detection using convolutional neural networks," *Sensors*, vol. 19, no. 7, p. 1644, 2019.

[10] T. Xu, Y. Zhou, and J. Zhu, "New advances and challenges of fall detection systems: a survey," *Applied Sciences*, vol. 8, no. 3, p. 418, 2018.

[11] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 5, pp. 611–622, 2011.

[12] L. Anishchenko, A. Zhuravlev, and M. Chizh, "Fall detection using multiple bioradars and convolutional neural networks," *Sensors*, vol. 19, no. 24, Article ID 5569, 2019.

[13] M. Yu, L. Gong, and S. Kollias, "Computer vision based fall detection by a convolutional neural network," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pp. 416–420, Glasgow, UK, November 2017.

[14] S. K. Yadav, A. Luthra, K. Tiwari, H. M. Pandey, S. A. Akbar, and S. A. Akbar, "ARFDNet: an efficient activity recognition & fall detection system using latent feature pooling," *Knowledge-Based Systems*, vol. 239, Article ID 107948, 2022.

[15] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, "UP-fall detection dataset: a multimodal approach," *Sensors*, vol. 19, no. 9, p. 1988, 2019.

[16] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis, "The mobiact dataset: recognition of activities of daily living using smartphones," *International Conference on Information and Communication Technologies for Ageing Well and E-Health*, vol. 2pp. 143–151, SciTePress, 2016.

[17] A. Sucerquia, J. López, and J. Vargas-Bonilla, "SisFall: a fall and movement dataset," *Sensors*, vol. 17, no. 12, pp. 198–1, 2017.

[18] D. Micucci, M. Mobilio, and P. Napoletano, "Unimib shar: a dataset for human activity recognition using acceleration data from smartphones," *Applied Sciences*, vol. 7, no. 10, p. 1101, 2017.

[19] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Analysis of public datasets for wearable fall detection systems," *Sensors*, vol. 17, no. 7, p. 1513, 2017.

[20] K. Soomro, A. R. Zamir, and M. Shah, "UCF101: a dataset of 101 human actions classes from videos in the wild," 2012, https://arxiv.org/abs/1212.0402.

[21] "UP fall detection dataset," https://sites.google.com/up.edu.mx/har-up/.

[22] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.

[23] P. Probst and A.-L. Boulesteix, "To tune or not to tune the number of trees in random forest," *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 6673–6690, 2017.

[24] H. Ramchoun, Y. Ghanou, E. Mohamed, and M. A. Janati Idrissi, "Multilayer perceptron: architecture optimization and training," *Multimedia and Artificial Intelligence*, vol. 4, 2016.

[25] S.-Mi Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using Convolutional Neural Network," in *Proceedings of the Ieee International Conference on Big Data and Smart Computing (Bigcomp)*, pp. 131–134, IEEE, Jeju, Korea (South), February 2017.

[26] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM Computing Surveys (CSUR)*, vol. 27, no. 3, pp. 433–466, 1995.

[27] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4489–4497, Santiago, Chile, December 2015.

[28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Ieee, Miami, FL, USA, June 2009.

[29] Mu. Li, T. Zhang, Y. Chen, J. Alexander, and Smola, "Efficient mini-batch training for stochastic optimization," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 661–670, New York, NY, USA, August 2014.

[30] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the International Conference on Machine Learning*, pp. 448–456, PMLR, Lille, France, July 2015.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, Honolulu, HI, USA, July 2017.

[33] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251–1258, Honolulu, HI, USA, July 2017.

[34] "Keras Homepage," https://keras.io/.

[35] "Scikit Library," https://scikit-learn.org/stable/.

[36] R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieva, and E. Moya-Albor, "A vision-based approach for fall detection using multiple cameras and convolutional neural networks: a case study using the UP-Fall detection dataset," *Computers in Biology and Medicine*, vol. 115, Article ID 103520, 2019.

Research Article

Eliminating Data Duplication in CQA Platforms Using Deep Neural Model

Seema Rani [ID], Avadhesh Kumar, and Naresh Kumar

School of Computing Science & Engineering Galgotias University, Greater Noida, Uttar Pradesh, India

Correspondence should be addressed to Seema Rani; seemanandal011@gmail.com

Primary research to detect duplicate question pairs within community-based question answering systems is based on datasets made of English questions only. This research put forward a solution to the problem of duplicate question detection by matching semantically identical questions in transliterated bilingual data. Deep learning has been implemented to analyze informal languages like Hinglish which is a bilingual mix of Hindi and English on Community Question Answering (CQA) platforms to identify duplicacy in questions. The proposed model works in two sequential modules. First module is a language transliteration module which converts input questions into a mono-language text. The next module takes the transliterated text where a hybrid deep learning model which is implemented using multiple layers is used to detect duplicate questions in the mono-lingual data. The similarity between the question pairs is done utilizing this hybrid model combining a Siamese neural network with identical capsule network as the subnetworks and a decision tree classifier. Manhattan distance function is used with the Siamese network for computing the similarity between questions. The proposed model has been validated on 150 pairs of questions which were scrapped from various social media platforms, such as Tripadvisor and Quora which achieves accuracy of 87.0885% and AUC-ROC value of 0.86.

1. Introduction

With the wireless access to the Internet becoming pervasive, recent years have seen an incessant advancement in Web usage with new user-centric applications and media-rich services. Social media has become an integral and indispensable part of our lives. Social networks make sharing of information, communication, and collaboration straightforward and opportune. The social media websites have grown significantly popular over the past decade as open-source platforms for information and knowledge sharing. Social media has become a primary source of the big data. Formally, big social data are the voluminous and varied data with high-velocity that is generated from technology-mediated social interactions and actions in online realm, and which can be scrapped and mined to comprehend social interactions and behavior. Undeniably, big data and social business intelligence are the key components for reaping full ROI.

Social media has been developing momentously and has changed the dynamics of the sharing and learning ecosystem. Social media news feeds and question answering sites are increasingly becoming popular valuable resource material for enriching knowledge base. The emerging role of social media in teaching-learning process cannot be ignored. Community Question Answering (CQA) [1] as a crowd-sourced service has emerged as a collective intelligence social system which facilitates participation of volunteers to express their knowledge and clear their uncertainties regarding some topics. It allows collaborative user participation where some users ask questions and the others provide answers to them. The number of participating users and the volume of questions and answers of different varieties characterize the popularity of CQAs.

Though these sites have high informational value while on the other hand in order to filter out the relevant information in the form of best answers, or answers by experts for reliable information and user satisfaction, it requires time

and labor. Some of the challenges associated with these platforms include ambiguity in user preference, duplicity in the questions, fuzziness, and imprecision among the huge and speckled answers [2]. The regular influx of new questions and getting answers to questions requires detecting semantic equivalence of questions. The problem of detecting semantic equivalence of multiple sentences is prevalent in natural language processing tasks. The problem further extends if there exists multilingual text or text using mash-up of varied languages like Hinglish.

Numerous CQA sites provide forum which contains answers to existing questions, and users are recommended to search the forum before posting their questions but it may not work as a same question may be rephrased in different ways and languages. There exist duplicate detection policies in some platforms, but these are limited to exact duplicate questions as well as near-exact duplicates. However, some of the challenges associated with analyzing questions for semantic equivalence detection are as follows:

(i) There are multiple ways to rephrase a same question

(ii) There may exist same solution to different questions

(iii) A question may be phrased in a mash-up of two languages

(iv) Merging of question threads is required after duplicate questions have been answered

(v) There exist volumes of historical questions

The first two cases define the issue where questions are in an asymmetrical domain. The examples in Figure 1 expound the first three cases.

Maximum number of primary studies on duplicate question detection deals with English-only CQAs. The cultural diversities, country-specific trending topics—hashtags on social media, and availability of native language keyboards add to the user-generated content in multiple languages contributing the linguistic challenges [3]. Linguistic anglicization of a language which is defined as to make a language sound or appear in English is a commonly observed phenomenon where one language/script word is transcribed into a source language so as to preserve source phonetics. For example, the question "Kya roz exercise krna zruri hai?" is a transliterated on the basis of pronunciation of bilingual mix of Hindi and English called Hinglish. The transliteration is not based on meaning. Instinctively, the question semantically becomes similar to the mono-lingual English question "Is daily exercise important?"

With recent upsurge in transliterated bilingual posts on CQAs which are restricted by their mono-lingual abilities and lack of collective part-of-speech taggers, it is essential to explore techniques that extend the cognitive capabilities of machines to interpret, comprehend, and learn features for semantic matching in multilingual (bilingual/trilingual) data. This research put forward a solution to resolve the problem of detection of semantic equivalent questions for duplicate detection in transliterated bilingual data. The proposed model examines the semantic similarity for input questions of varied languages using deep learning-based Siamese neural networks. In recent studies, evidences exist

where deep learning techniques are achieving advanced results on specific natural language processing-based problems as these possess hierarchical learning capabilities and are capable of generalizations [4, 5]. Siamese neural network is a deep learning-based technique comprising more than one identical subnetworks having same configuration (parameters and weights) [2, 6, 7]. The subnetworks of Siamese neural network are built on various categories of neural networks including convolutional neural networks and capsule networks. Semantic matching using Siamese neural network is performed by transforming the input sentences at each subnetwork into vectors and using a Manhattan distance function.

Thus, the proposed framework comprises two submodules first being the language transformation module and the second being the semantic matching module. As a first step, the input text containing questions is converted into a mono-language using language transliteration and translation tools. Further, a hybrid deep learning model is used for identification of duplicate questions in mono-language. Similarity detection between question pairs is carried out using a hybrid model using Siamese neural network combining with decision tree (DT) classifier. The Siamese neural network is built up of two identical capsule networks [8, 9] with same configuration, and Manhattan function is used as a distance measure for evaluating the similarity among questions. Further, DT classifier is used at the output layer which produces whether a question pair is duplicate or not. A dataset of 150 question pairs has been used to validate the proposed. The questions have been scrapped from multiple CQA forums like Tripadvisor and Quora.

The manuscript is structured as follows: Section 2 explores the literature review related to the domain of the study, Section 3 explains the proposed methodology, Section 4 presents the results, and Section 5 concludes the study.

2. Related Work

The growing number of Internet users in around interactive networking sites has increased researchers' interest to intelligently mine the user-generated online content. The utilization of machine learning and deep learning techniques for providing solutions to various problems of natural language processing tasks with user-generated online data has been evident to give accurate results. The potential of analyzing user-generated online content with intelligent data analysis techniques has provided solutions to various natural language processing problems which include sentiment analysis [4, 10], sarcasm detection [3, 11], rumor detection [12, 13], etc.

The large data generated by users each day on CQAs make it necessary to filter relevant data by discarding duplicate data. First step to achieve this is to analyze the data in order to detect duplicate questions by semantically matching the text. Multiple researchers have attempted to detect semantically equivalent sentences. Dey et al. [14] employed support vector machine (SVM) for identification of semantically equality in blogs using the SemEval-2015 dataset. Siamese neural network which encodes two input sentences
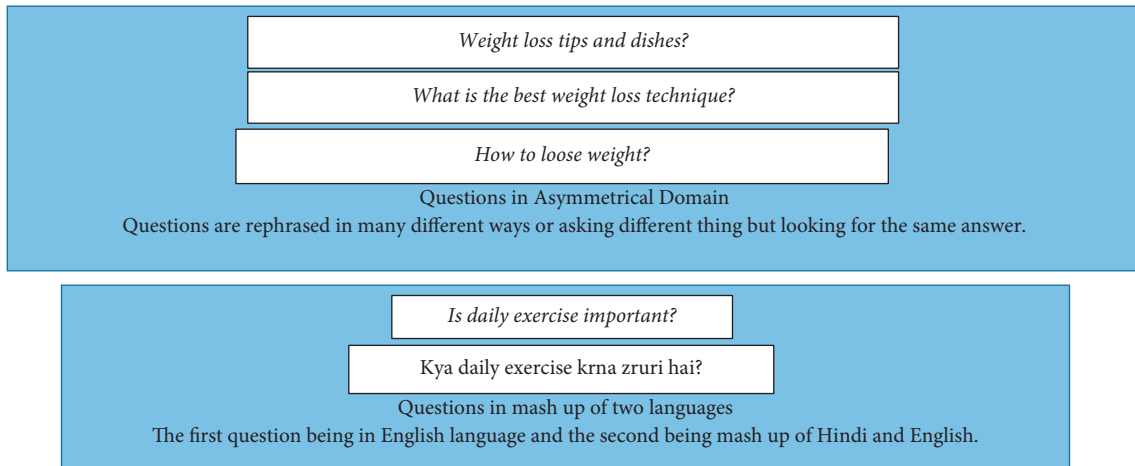
encoding those using same neural networks has been used recently to detect semantic equivalence of sentences and has shown accurate results [6, 7, 15]. In studies [16, 17], dataset of Quora has been analyzed for duplicate detection utilizing various connection networks. Rodrigues et al. [18] present performances of efficient experiments on detection of duplicity in questions using the established approaches and proposing a novel method. Chen et al. [19] attempt to solve duplicate question detection problem on Quora by a feature engineering and a neural-network-based model. Zhang et al. [20] propose a novel model named DupPredictor in order to recognize possible duplicates of an original question on stack overflow. In another study [21], the authors propose a model utilizing classification algorithms for duplicate question detection on stack overflow. They named their model as Dupe. Reproduction of DupPredictor as DupPredictorRep and Dupe as DupeRep has also been executed by Silva et al. [22] conducting advanced experiments. The use of Siamese neural network for resolving problem of duplicacy in CQA has also been seen in recent studies. Kumar [2] proposes a Siamese neural network model with attention-based bidirectional long short-term memory as subnetworks for detection of duplicate questions in medical CQAs.

The existing studies in the domain of matching of duplicate questions are done with English questions only as the current CQA forums provide facilities to post in one language only, mainly English only. Existing studies deal with question only in English language and deal with other language questions as noise. In a recent study [23], authors explored a code-mixed Q&A platform for Indian languages including Tamil and Hindi, their mash-up with English.

3. The Proposed Siamese Capsule Network with Manhattan Similarity

The proposed model comprises two modules. First one is the language transformation module, and second one is the semantic matching module (Figure 2).

The following sections describe the working of each module of the proposed framework in detail.

3.1. Language Transformation. The tokens of Hinglish are decoded in this module using transliteration into Hindi, and subsequently the Hindi words into English language are translated. "Transliteration is the process of converting a text in one language to text in another language on the basis of its pronunciation (maintaining it to greatest possible extent) without the change in grammar" [24]. Transliteration is different from translation in a manner that it is based on the pronunciation in the aimed language, while translation is based on the meaning in the target language (Figure 3). For instance, for the Hindi phrase "किस उम्र में बच्चों को इंटरनेट यूज़ करने देना चाहिए?" its translation in English would be "At what age should children be allowed to use the Internet?" and "Kis umar mein bachon ko Internet use karne dena chahiye?" is the transliterated Hindi. In this work, the Google Transliteration–Translator toolkit (Google Input Tools: https://www.google.co.in/inputtools/try/) has been used for the task of transliteration of code-mixed text into English in this module. The purpose of using a two-step transformation is imperative to capture the right textual interpretation. Using the Google Translate independently may not always output the correct meaning (Figure 3).

The correct interpretation for the example in Figure 3 would be "Should chocolate be eaten in depression?" and therefore, to improve the meaningfulness of transliterated text, the language transformation follows these steps (Figure 4):

(i) Word-level transliteration is done to convert Hinglish text to Hindi

(ii) Sentence-level translation is done to translate Hindi to English

This translated English question is then input to the Siamese network to match the semantics with the English-only input. The next section discusses the details about the semantic matching module.

3.2. Semantic Matching. Semantic matching of question pairs is done using a Siamese capsule network. Siamese neural network is made up of two neural networks of same
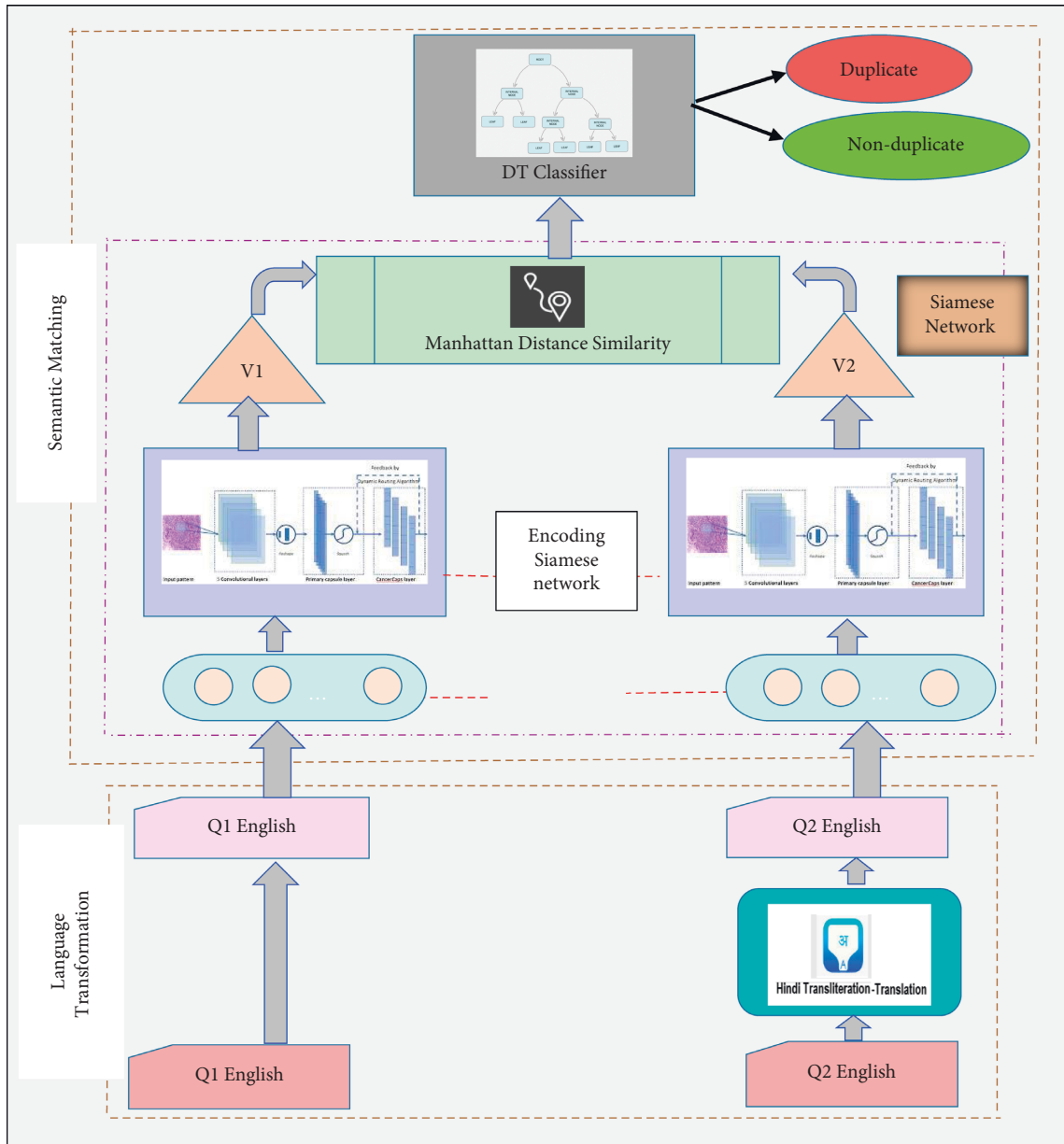
FIGURE 2: The proposed model.

configuration (parameters and weights) each taking different inputs. An energy function which compares feature representations of highest level on each side of the network by assessing a metric is used to join these two neural networks. Similarity among the two inputs (questions) is evaluated by transforming the inputs into representation vectors and using a distance function. We use Manhattan distance function for calculation of the similarity between all question pairs. The vector representations from the encoder are fed to the DT model together with their distance for binary classification.

3.2.1. Input Layer. This layer accepts questions as inputs into each subnetwork of the Siamese model. The question pair is further passed to the embedding layer.

3.2.2. Embedding Layer. In this study, GloVe [25] which is a count-based model is used for word embeddings in the embedding layer. In this layer, the input textual format is converted into vector table of words using GloVe. A real-valued vector of words is built as pretrained word embedding. Each question's embedding matrix is sent into the next layer.

3.2.3. Encoding Layer: Siamese Capsule Network. In this layer, the output of GloVe (word matrix) is converted into a one-dimensional feature vector. In this work, the encoding layer is employed using capsule network [8, 9]. The capsule network is composed of capsules which are groups of neurons. These groups are locally invariant groups with the capability of recognizing visual entities
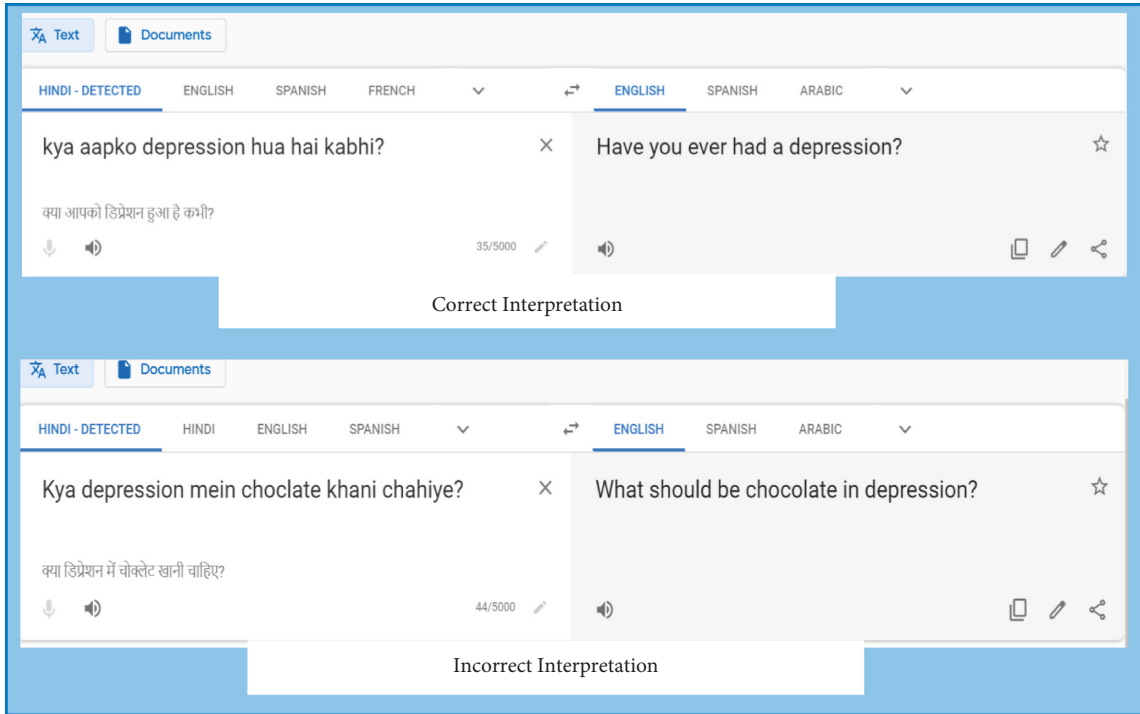
FIGURE 3: Sample independent use of Google translation.
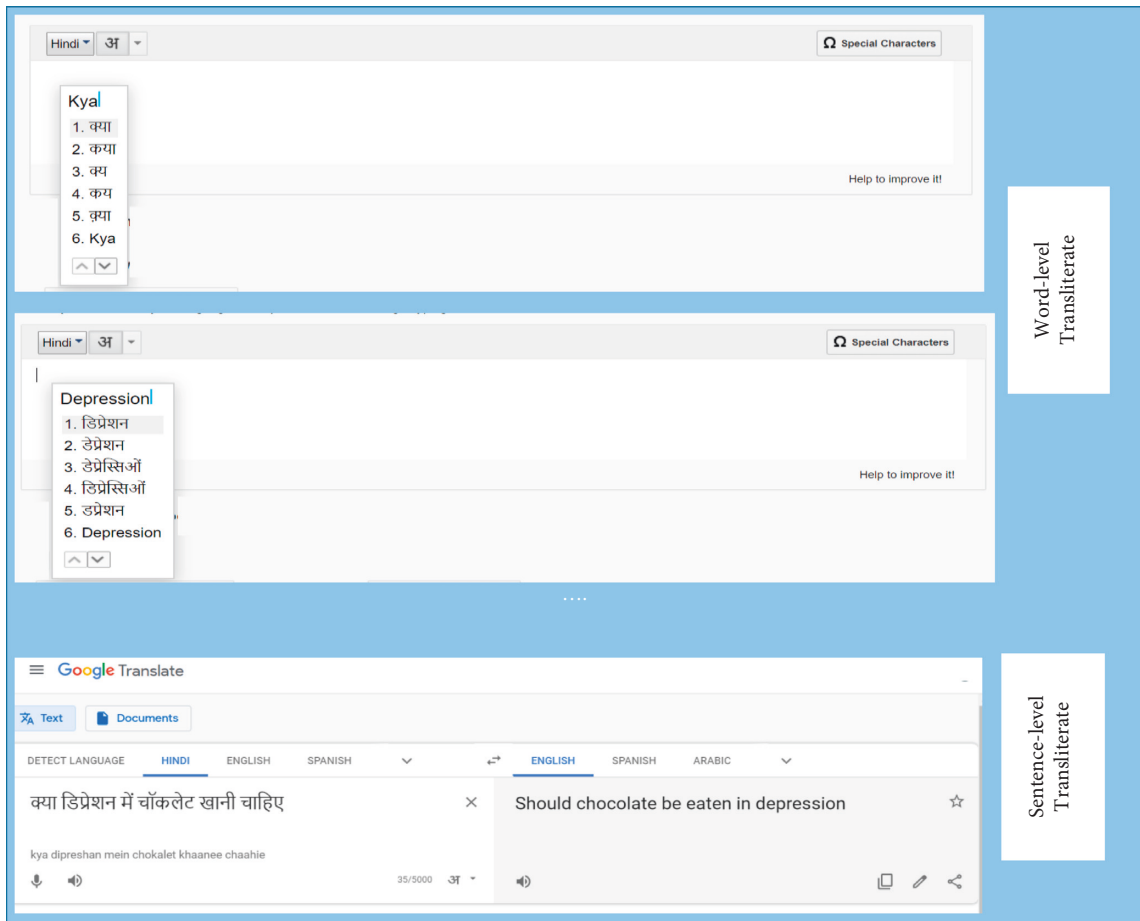
FIGURE 4: Transliteration: Hinglish to Hindi and translation: Hindi to English.

and encoding their features into vectors [26]. A capsule network comprises three layers including convolutional layer, primary capsule layer, and class capsule layer. Each layer provides a unique functionality given in following sections.

(1) Convolutional Layer. The vector representation of the text is provided to the filtration layer of the convolution having 128 filters of 8 size each, after applying a randomized function over the vectors. These filtered feature vectors are given as an input to a nonlinear function that acts as a linear function and uses stochastic gradient descent to train the model and to avoid the saturation problem of other activation functions. The rectified linear activation function has used the third layer of CNN, so named as ReLU activation layer. In each convolutional layer, the output from ith feature map, given by y_i^l, after the convolution operator is calculated as given in

$$y_i^l = b_i^l + \sum_{j=i}^{m} f_{i,j}^l * y_j^{l-1}, \tag{1}$$

where l is the layer having m filters, $f_{i,j}^l$ is the convolution filter, y_j^{l-1} is the output obtained from the previous layer, and b_i^l denotes the bias matrix.

(2) Primary Capsule Layer. The scalar inputs from the lower convolution layers are passed onto the first capsule layer. Here, these inputs are converted into vector outputs which are called capsules. Capsules intend to preserve the instantiate parameters including the syntactic or semantic representations of word and their local order. The features produced by the lower convolutional layers are further produced as combinations of features grouped into capsules in this layer, and a routing algorithm is implemented for communications with the upper layer. A capsule in the upper layer is activated if numerous lower layer capsules vote for that capsule. Results from lower level capsules are broadcasted to that activated upper level capsule. A cluster of capsules coming from lower layer is created at this activated capsule and a high probability of observation that an entity is present and a high-dimensional pose is produced as an output at the higher-level capsule.

For semantic matching of two sentences, it is difficult to say that two sentences are different because the order of the sentences is different. Hence, in this work we use the capsule network with routing algorithm as static routing algorithm.

Let u_i be the output of a single capsule, then the equation for static routing is as follows:

$$s_j = \sum_i W_{ij} u_i. \tag{2}$$

Here, capsule "i" is in the primary capsule layer which is the current layer, capsule "j" is a capsule in the higher level, and W_{ij} is translation matrix.

v_j is the output vector limited to [0, 1] as the result of applying squashing function to s_j as given below:

$$\mathbf{v}_j = \frac{\left|s_j\right|^2}{1 + \left|s_j\right|^2} \frac{s_j}{\left|s_j\right|}. \tag{3}$$

(3) Class Capsule Layer. This layer is the capsule network's final capsule layer which takes inputs from the lower layer through a static routing algorithm, also called the routing by agreement.

(4) Similarity Measure. This study utilizes the exponential function on the negative value of distance measured by Manhattan distance function to find the similarity distance between the vector representations. This is represented in

$$\exp\left(-\|v1 - v2\|\right) \in [0, 1], \tag{4}$$

where v_1 and v_2 represent the outputs of subnetworks.

3.2.4. Output Layer. Decision tree (DT) classifier is implemented at the output layer to make the final prediction. DT is a tree-based model which comprises a root, branches, and leaf nodes where the dataset information is broken down into smaller subsets till an associated decision tree is built. The input attributes are allied with the nonleaf nodes, and the classes are represented by the leaf nodes. The output of Siamese network is sent to DT classifier which models the interaction between the two questions. The output of Siamese encoder is fed to DT generating a hybrid Siamese neural network and DT classifier which models the interaction between the two questions. The concatenation of the representations of questions with their distance, $c = [f(q1); f(q2); d(q1, q2)]$, is given as input to DT, and the probability of a match between them is produced as output.

4. Results

Section 4.1 discusses the dataset used in this study to validate the proposed model, and Section 4.2 discusses performance of the proposed hybrid model. Section 4.3 describes the performance analysis of various question types, and Section 4.4 compares the Manhattan distance measure with various other distance measures. Lastly, Section 4.5 shows the comparison of the accuracy of the proposed hybrid model with different variants of Siamese subnetworks.

4.1. Dataset. The data are prepared by extracting 150 pairs of questions from CQA forums including Tripadvisor and Quora. It consists of pairs of questions each having one question in English, and one is in Hinglish script. The dataset is labeled with "duplicate" and "nonduplicate" for each question pair. Table 1 presents a sample snapshot of the data.

4.2. Performance of Proposed Framework. The performance of the proposed Manhattan hybrid Siamese neural network model with DT classifier for similarity matching is evaluated in terms of accuracy and AUC-ROC. "Accuracy is defined as the proportion of correct predictions to the total number of

TABLE 1: Snapshot of dataset.

English question	Hinglish question	Is duplicate?
Which animal would you like to pet	Aapka favorite jaanvar kya hai	N
Are you satisfied with your life	Kya aap jeevan m aur paise kamana chahte hai	N
What are the benefits of eating mango	Mango khaane k kya faayde hai	Y
How to follow your passion	Apne passion ko kaise sakar kare	Y
How to lose belly fat	Kamar k aas paas ka fat kaise ghataye	Y
Should we drink green tea	Kaunse brand ki green tea peeni chaiye	N
How much does a road trip from LA to Vegas cost	Las Vegas se LA jaane ke liye kya options hain	N
Is daily exercise important	Kya daily exercise krna zruri hai	Y
How to open WhatsApp Web on desktop	Desktop pe WhatsApp kaise use karte sakte hain	Y
Which places to visit in Mumbai	Mumbai mein best beach kaunsa hai	N

predictions." "ROC curve (receiver operating characteristic curve) is a graph measuring the performance of a classification model by plotting true-positive rate against false-positive rate." "AUC stands for area under the ROC curve which measures the entire two-dimensional area underneath the entire ROC curve." Optimal selection of parameters is imperative to achieve superlative performance results. The hyper-parameters are tuned using the validation data in order to obtain accurate results. GloVe is implemented with dimensions of 300 and a batch size of 60 and 0.5 drop-out.

The confusion matrix and ROC graph plots for the same are given (Figures 5 and 6). The confusion matrix is a plot of true positive, false positive, false negative, and true negative. The model attains an accuracy of 87.0885%. The results show that it is predicting accurate results in identifying correct and incorrect answers. The value of the precision column indicates low false-positive rate. It does not identify wrong answers as correct answers. Higher recall value suggests that our system makes less mistakes in classifying the answers of a question. It can correctly predict that is the answer right or wrong. Because the recall value is high, there are fewer forecasts when the actual answer is true but the system predicted incorrectly.

Finally, $F1$ score is basically the average of two quantities that is the precision and the recall in which we consider the false positives and the false negatives. Higher $F1$ score establishes that our system makes less mistakes.

4.3. Performance on Different Question Types.
The accuracy of the proposed hybrid framework has been evaluated for various question types such as "what," "who," "why," "when," "where," "which," "are," "is," "should," and "how" using the exact match measure which is represented in Figure 7.

It can be seen that the best predictions can be made for questions of type "when." The next best predictions are made for the questions beginning with "who" and "how," whereas least accuracy was achieved with the "why"-type questions.

4.4. Comparison of Various Similarity Measures.
The performance of the model using the Manhattan distance has been compared with the models using Euclidean and cosine distances as similarity measures. The results are shown in Figure 8.
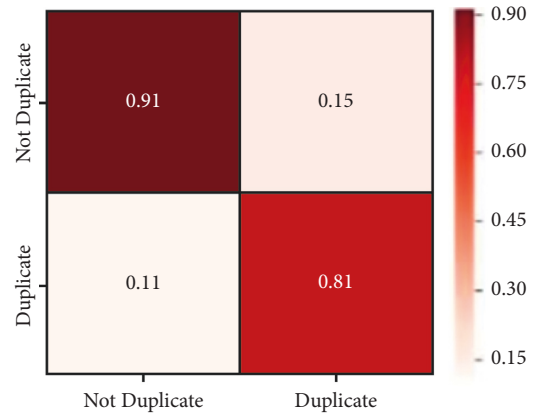
FIGURE 5: Confusion matrix.
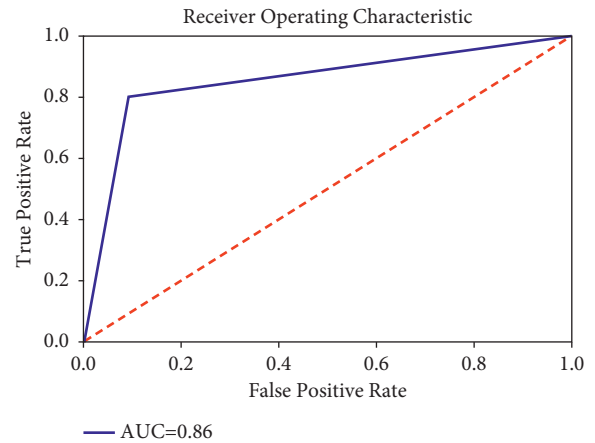
FIGURE 6: ROC curve.

The model using Manhattan distance performs best, and the model using cosine distance gives the worst performance.

4.5. Assessment of Other Subnetworks.
The accuracy of the proposed hybrid model has also been compared with Siamese models having different neural network architectures at the subnetworks. The other models which have been evaluated for the detection of duplicate question pairs in transliterated bilingual data are as follows:

Performance based on Types of questions
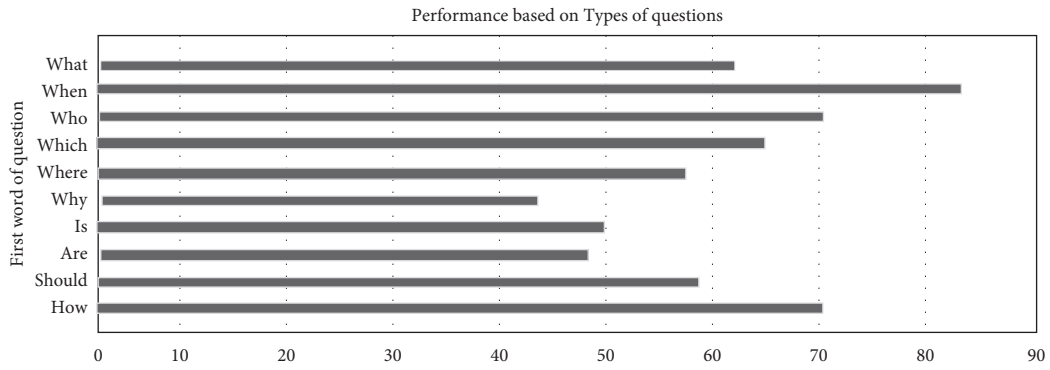
Figure 7: Accuracy obtained on different question types.

	Manhattan Distance	Euclidean Distance	Cosine Distance
Accuracy	87.08	80.65	76.42
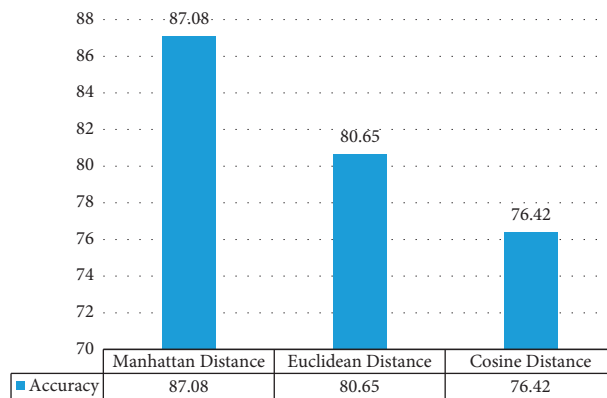
Figure 8: Comparative performance using varied distance functions.
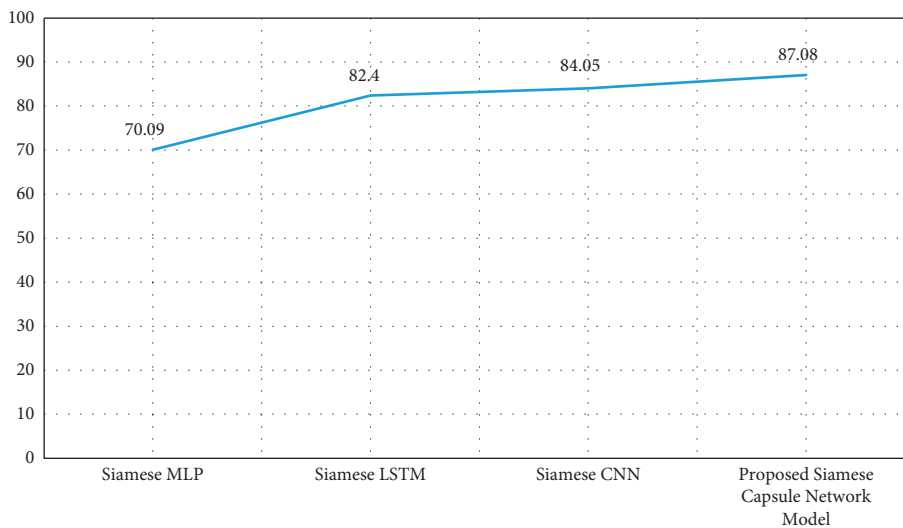
Figure 9: Performance of various subnetworks.

(i) Siamese MLP network

(ii) Siamese LSTM network

(iii) Siamese CNN

Manhattan distance for similarity measure and the same dataset has been used in all the models. The comparative performances are given in Figure 9 where the y-axis denotes the accuracy percentage plotted for each model on the x-axis.

As seen in the figure, the proposed methodology utilizing hybrid Siamese capsule network with DT classifier and Manhattan similarity outperforms the other Siamese neural network models' semantic similarity matching in transliterated bilingual data.

5. Conclusion

The large volume big data generated through the communications through social networks makes it imperative analysis for data-driven decision making. But adversarial IR (targeting, probing, spam traps), multilingual, and duplicate content are the key issues that challenge the performance of a typical CQA. This study proposes a hybrid model combining a Siamese neural network with capsule network as a subnetwork and a decision tree classifier to detect duplicate pairs of questions. For the problem of semantic question matching for duplicate detection in transliterated bilingual data on CQA, the input questions are transformed into a single language using language transliteration and translation tools. Manhattan distance similarity measure is used with the deep learning model for similarity match between the question pairs. The proposed hybrid model achieves an accuracy of 87.0885%. The limitation of the work includes lack of benchmark datasets. In future, we intend to build a language-independent framework for identifying semantic similarity which could comprehend duplicates in bilingual or trilingual data. The model can also be trained and tested for visual question answering. Further studies on types of mash-up including code-mix and code-switch input types can also be done.

Data Availability

Data will be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Kumar and D. Kumar Gupta, "Empirical analysis of supervised machine learning techniques for expert mining in CQA system," *International Journal of Information Systems & Management Science*, vol. 2, no. 2, 2019.

[2] A. Kumar, "Using cognition to resolve duplicacy issues in socially connected healthcare for smart cities," *Computer Communications*, vol. 152, pp. 272–281, 2020.

[3] D. Jain, A. Kumar, and G. Garg, "Sarcasm Detection in Mash-Up Language Using Soft-Attention Based Bi-directional LSTM and Feature-Rich CNN," *Applied Soft Computing*, vol. 91, Article ID 106198, 2020.

[4] A. Kumar, K. Srinivasan, W. H. Cheng, and A. Y. Zomaya, "Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data," *Information Processing & Management*, vol. 57, no. 1, Article ID 102141, 2020.

[5] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing [review article]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[6] D. Bogdanova, Cicero dos Santos, L. Barbosa, and B. Zadrozny, "Detecting semantically equivalent questions in online user forums," in *Proceedings of the 19th Conference on Computational Natural Language Learning*, pp. 123–131, Beijing, China, July, 2015.

[7] J. Bromley, J. W. Bentz, L. Bottou et al., "Signature verification using a "siamese" time delay neural network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 4, pp. 669–688, 1993.

[8] W. Zhao, J. Ye, M. Yang, Z. Lei, S. Zhang, and Z. Zhao, "Investigating Capsule Networks with Dynamic Routing for Text Classification," 2018, https://www.arxiv-vanity.com/papers/1804.00538/.

[9] J. Kim, S. Jang, E. Park, and S. Choi, "Text classification using capsules," *Neurocomputing*, vol. 376, pp. 214–221, 2020.

[10] A. Kumar and A. Jaiswal, "A deep swarm-optimized model for leveraging industrial data analytics in cognitive manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 17, 2020.

[11] A. Kumar and G. Garg, "Empirical study of shallow and deep learning models for sarcasm detection using context in benchmark datasets," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–16, 2019.

[12] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar, and M. Abdel-Basset, "Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network," *IEEE Access*, vol. 7, Article ID 23319-23328, 2019.

[13] S. R. Sangwan and M. P. S. Bhatia, "D-BullyRumbler: A Safety Rumble Strip to Resolve Online Denigration Bullying Using a Hybrid Filter-Wrapper Approach," *Multimedia Systems*, pp. 1–17, 2020.

[14] K. Dey, R. Shrivastava, and S. Kaushik, "A paraphrase and semantic similarity detection system for user generated short-text content on microblogs," *COLING*, vol. 16, pp. 2880–2890, 2016.

[15] A. Sanborn and J. Skryzalin, *Deep Learning for Semantic Similarity. CS224d: Deep Learning for Natural Language Processing Stanford*, Stanford University, CA, USA, 2015.

[16] G. Wang, K. Gill, M. Mohanlal, H. Zheng, and B. Y. Zhao, "Wisdom in the social crowd: an analysis of quora," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 1341–1352, ACM, Rio de Janeiro, Brazil, 2013 May.

[17] C. Saedi, J. Rodrigues, J. Silva, and V. Maraev, "Learning profiles in duplicate question detection," in *Proceedings of the 2017 IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 544–550, San Diego, CA, USA, 2017 August.

[18] J. A. Rodrigues, C. Saedi, V. Maraev, J. Silva, and A. Branco, "Ways of asking and replying in duplicate question detection," in *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (SEM 2017)*, pp. 262–270, Vancouver, Canada, August, 2017.

[19] Z. Chen, H. Zhang, X. Zhang, and L. Zhao, *Quora Question Pairs*, 2018.

[20] Y. Zhang, D. Lo, X. Xia, and J. L. Sun, "Multi-factor duplicate question detection in stack overflow," *Journal of Computer Science and Technology*, vol. 30, no. 5, pp. 981–997, 2015.

[21] M. Ahasanuzzaman, M. Asaduzzaman, C. K. Roy, and K. A. Schneider, "Mining duplicate questions in stack overflow," in *Proceedings of the 13th International Conference on Mining Software Repositories*, pp. 402–412, ACM, Austin, Texas, USA, 2016, May.

[22] R. F. Silva, K. Paixão, and M. de Almeida Maia, "Duplicate question detection in stack overflow: a reproducibility study," in *Proceedings of the 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pp. 572–581, IEEE, Campobasso, Italy, 2018, March.

[23] K. R. Chandu, M. Chinnakotla, A. W. Black, and M. Shrivastava, "Webshodh: a code mixed factoid question answering system for web," in *Proceedings of the International Conference of the Cross-Language Evaluation Forum for European Languages*, pp. 104–111, Springer, Thessaloniki, Greece, 2017, September.

[24] K. Knight and J. Graehl, "Machine transliteration," *Computational Linguistics*, vol. 24, no. 4, pp. 599–612, 1998.

[25] J. Pennington, R. Socher, and M. C. Glove, "Global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543, (EMNLP), Doha, Qatar, October, 2014.

[26] E. Xi, S. Bing, and Y. Jin, "Capsule Network Performance on Complex Data," 2017, https://arxiv.org/abs/1712.03480.

Retraction

Retracted: On the Design of Secured and Reliable Dynamic Access Control Scheme of Patient E-Healthcare Records in Cloud Environment

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] K. Zala, H. K. Thakkar, R. Jadeja et al., "On the Design of Secured and Reliable Dynamic Access Control Scheme of Patient E-Healthcare Records in Cloud Environment," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3804553, 19 pages, 2022.

Hindawi

Research Article

On the Design of Secured and Reliable Dynamic Access Control Scheme of Patient E-Healthcare Records in Cloud Environment

Kirtirajsinh Zala [iD],[1] **Hiren Kumar Thakkar** [iD],[2] **Rajendrasinh Jadeja** [iD],[3] **Neel H. Dholakia**,[1] **Ketan Kotecha** [iD],[4] **Deepak Kumar Jain**,[5] **and Madhu Shukla**[1]

[1]*Department of Computer Engineering, Marwadi University, Rajkot 360006, Gujarat, India*
[2]*Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar 382007, Gujarat, India*
[3]*Faculty of Technology, Marwadi University, Rajkot 360006, Gujarat, India*
[4]*Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed) University, Pune, India*
[5]*Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China*

Correspondence should be addressed to Hiren Kumar Thakkar; iamhiren@gmail.com and Ketan Kotecha; head@scaai.siu.edu.in

Traditional healthcare services have changed into modern ones in which doctors can diagnose patients from a distance. All stakeholders, including patients, ward boy, life insurance agents, physicians, and others, have easy access to patients' medical records due to cloud computing. The cloud's services are very cost-effective and scalable, and provide various mobile access options for a patient's electronic health records (EHRs). EHR privacy and security are critical concerns despite the many benefits of the cloud. Patient health information is extremely sensitive and important, and sending it over an unencrypted wireless media raises a number of security hazards. This study suggests an innovative and secure access system for cloud-based electronic healthcare services storing patient health records in a third-party cloud service provider. The research considers the remote healthcare requirements for maintaining patient information integrity, confidentiality, and security. There will be fewer attacks on e-healthcare records now that stakeholders will have a safe interface and data on the cloud will not be accessible to them. End-to-end encryption is ensured by using multiple keys generated by the key conclusion function (KCF), and access to cloud services is granted based on a person's identity and the relationship between the parties involved, which protects their personal information that is the methodology used in the proposed scheme. The proposed scheme is best suited for cloud-based e-healthcare services because of its simplicity and robustness. Using different Amazon EC2 hosting options, we examine how well our cloud-based web application service works when the number of requests linearly increases. The performance of our web application service that runs in the cloud is based on how many requests it can handle per second while keeping its response time constant. The proposed secure access scheme for cloud-based web applications was compared to the Ethereum blockchain platform, which uses internet of things (IoT) devices in terms of execution time, throughput, and latency.

1. Introduction

In the recent past, data volumes are exponentially increased each passing day due to affordable and easy access to internet-enabled connected devices [1, 2]. The healthcare systems are under constant strain to deal with a high volume of healthcare data with the expectation to predict the results in a reasonable time duration. Systems and algorithms are developed to deal with the high velocity of the healthcare data for efficient prediction. However, most of such systems are subject to data security risk and face attacks such as denial of service (Dos), snooping, and traffic analysis [3]. In addition, the recent coronavirus pandemic has forced the hospitals to run at overcapacity making it difficult to keep

patient records secure. The conventional method of manual handwritten paper-based information storage is not viable and makes it difficult to efficiently retrieve the data. Moreover, traditional methods make it challenging to find a patient's individual medical and data that are subject to be lost or destroyed [4]. Using blockchain and Interplanetary File System (IPFS), the author [5] proposes an architecture that aims to offer quicker retrieval and consistent personal health record availability. The findings demonstrate that an ideal node is chosen among all potential adjacent nodes in each iteration.

The internet of things (IoT) technology enables the storage of health records in digital format [6, 7]. Because e-healthcare security includes the patient's confidential health information, it is more essential to be in a secured form. To carry out attacks, attackers can take advantage of the vulnerabilities in open wireless channels [8–11]. These attacks have the potential to harm the e-healthcare system in various ways. An example of this would be a patient who has been treated and then sent home from a hospital in a place other than his own. The patient later becomes ill and is hospitalized nearby, but it does not have all of the data or the records from his previous hospitalization. Due to a lack of information, his therapy may be delayed or fatal. However, if the patient's data are already stored on cloud-connected devices, retrieving the patient's data takes only seconds, allowing the new hospital personnel to begin treatment earliest possible [12].

Using high-security cryptography techniques, the healthcare department can keep encrypted data in the cloud, limiting access to only those who have been granted permission. The cloud comprises servers that run various software and databases and can be accessed via the internet. Cloud servers can be found worldwide to store and access data from any place, anywhere. Because of cloud computing, the healthcare departments and insurance companies do not need to make use of any physical servers or any software programs [13]. The capacity to quickly and securely transfer massive volumes of data, such as patient medical records, is one of the many advantages of cloud computing [14]. Healthcare providers should use digital solutions in hospitals to ensure that their infrastructure is well managed and that they have adequate opportunities to engage themselves with IT service providers [15, 16]. Cost-effectiveness, collaborative resource sharing, scalability, and agility enhancement are some of the other advantages of mobile and cloud computing [17].

Electronic health data (EHD), personal health record (PHR), and electronic medical record (EMR) are all types of digital medical records [18]. Healthcare professionals keep EMR and EHR, whereas the patients or their relatives own PHR. Man-in-the-middle attack, denial of service (DOS), eavesdropping, and so on are all threats to wireless communication among physicians and patients, and between the cloud and the systems.

Healthcare providers and patients can benefit from the cloud's ability to store, process, and update information without spending a lot of money and its ability to make things more efficient and of better quality. Since this information is stored on more than one server, it can be easily accessed from many different places. E-health systems guarantee on-demand, quick, and consistent access to health records, and fewer medical errors and higher-quality treatment. But they also leave patient privacy open to misuse of EHR data and improper authorization. So, security and privacy are very important when multiple people share or look at patient data. Figure 1 gives an overview of the architecture of e-health.

Although cloud services offer enormous benefits, they still face numerous security risks. The cloud service provider, for example, has a vast amount of data that users do not know about [19]. A lack of transparency makes it impossible to know how data are handled and where. Because of this, it is harder to have faith in the service provider, and data loss may result. Figure 1 shows the vulnerability of untrusted servers to assaults from both internal and external adversaries because they lack privacy-preserving safeguards.

The CP-ABE systems with multiple authorities and threshold secret sharing (t, n) can be integrated; this research [20] presents an improved security and performance for public cloud storage by addressing the single-point bottleneck problem, which enhances both security and performance. Using an auditing technique, a solitary bottleneck issue for the most existent CP-ABE systems can be alleviated, according to the author [21]. E-health cannot benefit from these advanced access control schemes [20, 21] despite the fact that the central authority and many attribute authorities cannot ensure protection from insider attacks, even though they are advanced access control schemes with excellent security measures. Water-based CP-ABE (cyphertext-policy attribute-based encryption) system deniable on attribute-based encryption (ABE) technique is a particular encryption approach that allows cloud storage providers to build fake user identities using preserved cyphertext in order to safeguard data from external attackers [22].

By encrypting data belonging to many patients who share the same access policy mentioned by [23], the author's system offers multiprivileged access control for personal health records (PHRs). For disease prediction, author [24] uses the single-layer perceptron learning algorithm. Using encrypted prediction models developed by this model, the cloud leverages encrypted symptom data supplied by patients to diagnose their illness without jeopardizing patient privacy. Health records cannot benefit from these procedures because of their high level of data privacy and computational complexity and scalability concerns [23, 24]. A CP-ABE with the disguised access control mechanism policy and permitted access control was proposed in another study, [25].

A new identity-based encryption (IBE) method based on revocable storage presented by the author [26] protects cyphertext in both the forward and backward directions. There is currently no dynamic user management among the most secure provenance cloud storage systems, which results in considerable performance overhead and poor access control. An attribute-based, provenance-assured cloud storage system is presented in this paper [27], which offers an answer to the issue. As efficient as ABE schemes are, it is
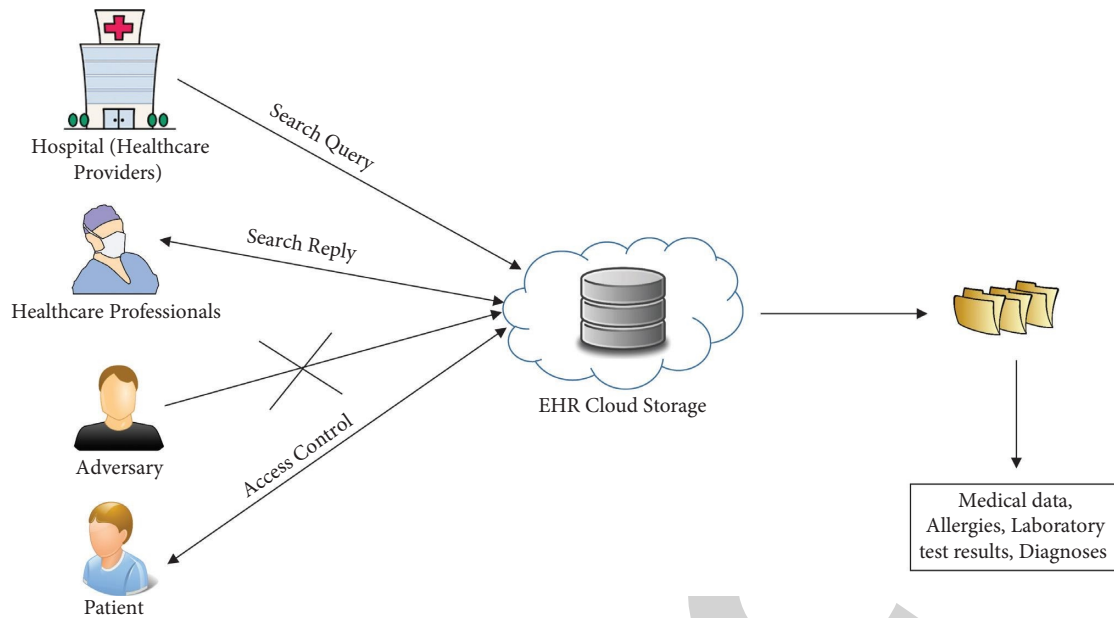
FIGURE 1: Cloud-based electronic health data architecture.

still not possible to implement them on EHRs. This is despite the fact that ABE schemes provide perfectly alright, well-formed access to health records [25, 27]. There are several reasons for this, including high computing costs, the difficulty of managing keys, and the inability to effectively administer access control regulations [22]. A policy that offers approved access control with a constant key length increases as the number of attributes in the access structure increases [25].

The e-healthcare technology should also protect patient's data, even if the healthcare department argues that personnel are responsible for doing so. Our study suggests a system in which only authorized personnel have access to patient information. The doctor has read and wrote access to the data, whereas others can only read it but not edit it. To guarantee the protection of sensitive data from beginning to end, the suggested technique outlines how an administrator generates subkeys with help of the master key. Confidentiality, authenticity, and protection from known critical attacks were all included in our study of healthcare system security [28, 29].

1.1. Motivation and Contribution.
There are not enough privacy protection systems in place to guarantee complete safety in the cloud for e-health data. Health records kept on cloud servers face the greatest risk from insider assaults, such as those by database administrators or key managers, rather than outsiders as it is contrary to popular opinion. Furthermore, cloud access could expose potentially sensitive information to malicious users if not secured. Risks to patient's lives increase as a result of its medical usage. The malicious user may also use the information to harm the hospital's reputation. Future solutions to these cyberattacks and criminal individuals may include the adoption of secure access control technologies. However, due to asset user

devices and insecure wireless channels, it is challenging to build such protocols. As a result, lightweight security mechanisms with excessive reliability should be developed to protect sensitive patient data over the third-party cloud service provider.

(i) We created a single cloud-based web application used to store and provide secure access to e-health records from anywhere. The web application is hosted on the AWS (Amazon Web Services) cloud platform by launching multiple instances. We also investigated the performance of our web application in terms of specific requests per second. Performance is evaluated by looking at resource provisioning, CPU utilization, throughput, and response time.

(ii) We compared the performance evolution of a secure access scheme for a cloud-based web application for patient health records with the Ethereum blockchain platform using IoT devices.

(iii) We propose a bold and secure method of entry for e-healthcare services in the cloud-based scheme.

(iv) In accordance with the recommended security protocols, only legitimate users can use cloud services, which confirm the user's identification in cloud-based e-health application (patient, doctor, and ward boy).

(v) The proposed protocol protects networks from message modification, replay, and man-in-the-middle attacks while ensuring data confidentiality, message freshness, and other security features.

The rest of this paper is organized as follows. Section 2 discusses the literature review, Section 3 describes the system design and adversary prototype, Section 4 explains the proposed scheme for secure access, Section 5 presents the

experimental setup with results and discussion, Section 6 discusses experimental design, and Section 7 ends with the conclusion.

2. Literature Review

According to a study [30], a new web-based solution that allows doctors, ward boys, and pharmacists to access patients' medical records has been developed. It stores the patient's information on the local cloud. The data can be accessed and updated from a distance. To collaborate on treatments, records must be provided to other doctors. The disadvantage in the method is it prevents patients from accessing their medical records.

The researchers of [31] proposed a healthcare framework based on the cloud for effective collaboration among caregivers and the healthcare providers, which might fully change hospital's handwritten record systems. The records can be accessed by healthcare providers and patients from any location utilizing the system. The authorization management service, part of the framework, could only be accessed by a genuine healthcare practitioner or patient. In contrast to patients, healthcare professionals are allowed to write, read, and alter the data. There are two portions to a patient's health record; one is locally stored in a healthcare facility, while another is stored on a cloud database server. The biggest issue with this approach is that the cloud server stores all data when a hospital or healthcare facility does not have a local EHR system.

In paper [32], experts examined, studied, and evaluated several papers and found that several issues need to be addressed to preserve e-healthcare records (EHR). EHR security, EHR cloud architecture, and EHR privacy are only a few. The authors also state that there are still many studies to be performed in the field of EHR security. Author [33] proposes a solution to the problem of managing user access control to a complex universe of user data while maintaining confidentiality while storing medical records in the cloud. Author [34] described an authentication method for an EHR system with a hybrid cloud structure, allowing us to handle different types of users with varying access privileges.

The authors [35] propose a simple and effective method for securing healthcare institution collaboration. They present secure multiparty computation (SMC) methods to assure compliance with data protection regulations. When outsourcing computations to the public cloud, the authors employ the Paillier scheme to safeguard medical data from unauthorized access. Another advantage of this technique is its ability to execute arithmetic operations on encrypted data without access to the original data. Using Shamir's secret sharing, the author proposes a novel cloud storage system for EHRs that ensures data privacy in its entirety. In this system, a healthcare facility divides an EHR into multiple segments, which are then distributed to numerous cloud servers. When retrieving EHRs, the healthcare facility extracts segments from partial cloud servers and reconstructs EHRs [36].

The suggested method listed in paper [37] is based on the FHE (fully homomorphic encryption) algorithm with key delegation to ensure data privacy, authentication, integrity,

and availability in a hierarchical structure with multiple levels. This will give the healthcare provider the freedom to use or not use any access rule in any order, which is especially important in a medical research setting. Still, there is more work to be done to make FHE really useful. Flow-enabled distributed mobility anchoring (FDMA) was proposed by the author to reduce signaling overhead cost (SOC) and packet tunneling cost (PTC) [38].

The researcher in paper [39] presents the experimental investigation of cryptographic algorithms to classify encryption algorithm types such as symmetric and asymmetric. It provides a comprehensive analysis of advanced encryption standards (AES), data encryption standards (DES), 3DES, RSA, and Blowfish in terms of timing complexity, file size, encryption, and decryption performance. The speed of encryption and decryption of the selected encryption algorithms has been evaluated utilizing a simulation-based methodology. The research [40] proposes using the distributed fog computing architecture that uses the elliptic curve Diffie–Hellman ephemeral (ECDHE) key exchange algorithm with the preshared key (PSK) as an authentication method that is both lightweight and safe. As an alternative to the static PSK technique, the ephemeral preshared key used by the ECDHE-PSK authentication system provides perfect forward secrecy (PFS). Literature [41] provides a lightweight, reliable encryption scheme for healthcare image encryption.

The purpose of paper [42] is to discuss data security and authentication in the healthcare industry. Author [43] has proposed a blockchain-based, public-key cryptography-based secure framework. The authors of [43] suggest a model in which medical pictures from the Digital Imaging and Communications in Medicine (DICOM) standard, which contains data on disease and may be applied in real time to the healthcare system, are shared. Due to the blockchain-based decentralized storage model, the framework keeps the immutability, privacy, and availability of information. They also discussed how peers inside the blockchain network could access information via consensus, which they explained in detail. To solve the problem of scalability, the author [44] suggests an improved version of the Bell–LaPadula model and divides peers and transactions into groups with different levels of clearance and security. Due to the clearance level, the peers do not have to keep track of every transaction made. Using smart contracts, author set up dynamic access control policies in the network to keep data safe. Author uses a blockchain-based healthcare network to test their model [44].

Today, this large amount of medical data made by IoMT (internet of medical things) is kept in a centralized storage system. However, centralizing all of a patient's private information raises questions about security and privacy. To deal with these problems, author suggests a consortium blockchain network that can handle smart contracts. In the initial stage of patient and medical device authentication, the author built an integrated interplanetary file system cluster node that is an interplanetary file system using smart contracts. In order to safely transfer device-generated data throughout the consortium blockchain, it is proposed that the same cluster layer has been used as a distributed data storage layer [45].

After the study of various comparative literature review of cryptographic and noncryptographic methods, decentralized EHR of patient for legitimate users with centralized management of patient data security is one of the most challenging problems when using cloud systems. It is one of the main reasons why many organizations in e-healthcare avoid using cloud services. The issue is how cloud systems for integrating decentralized information systems must be constructed in terms of technology and organization so that cloud user privacy laws may be guaranteed. Our paper proposes a secure scheme implemented in a cloud-based e-health application in which only authorized personnel can access patient data. Those outside of the doctor's office can only view the data, but they are unable to change or add to the information. The proposed scheme implemented in cloud-based e-health application is designed for decentralized EHR with centralized management. Moreover, in the industrial IoT, there is a growing demand for a privacy-preserving secured framework. In [46], a deep blockchain-based trustworthy privacy-preserving secured framework in the industrial internet of things systems is proposed.

Patient EHRs are decentralized in terms of providing selected rights to third parties (doctors, ward boys, relatives, and hospitals) with access to critical resources (information, applications, EHR, and reports) to enable more effective operations in application. The "Storage Cloud" cloud computing technology is suitable for the technological implementation because it facilitates the simple incorporation of user data storage into the system. Data administration and authoritative authority are delegated to a centralized service provider that does not store user data. The user acquires a greater degree of responsibility and authority as a result of the breakdown of trust. A guarantee of data availability is essential for the provision of high-quality services. Applications containing EHR are run in a decentralized manner but are centrally managed.

2.1. Comparative Analysis of Proposed Scheme. Table 1 provides a clear comparison of the security features of the existing and new protocols. The suggested approach achieves all of the significant security features, for example, secrecy, integrity, authorization, anonymity, and authentication, as shown in the first row of Table 1. On the other hand, traditional approaches cannot achieve all of them, as evidenced by all except the first row of the table. The failure of any standard system to accomplish all basic security features indicates potential weaknesses and high attack possibilities.

Table 2 compares and contrasts the different protocols on e-healthcare security. Table 2 details that many risks attackers can use to launch cyberattacks against different medical devices. Table 2 includes the level of complexity required to successfully hack into medical devices and the level of awareness that stakeholders must have in order to successfully hack into them. The strategies in some of the research paper presented here in literature review do not provide perfect security regarding authenticity, anonymity of one's identity, communication integrity, and secrecy. Because traditional schemes lack these security features, they

TABLE 1: Analyzing protocols based on security features.

Scheme	J_1	J_2	J_3	J_4	J_5
J_S	Yes	Yes	R	Yes	Yes
[12]	No	No	R	No	No
[47]	No	Yes	W_O	No	Yes
[14]	No	Yes	R	Yes	Yes
[48]	No	Yes	W_O	No	Yes
[49]	No	Yes	R	Yes	Yes
[50]	No	Yes	R	Yes	Yes
[28]	No	Yes	W_O	Yes	No
[50]	No	Yes	W_O	No	No
[51]	No	Yes	W_O	Yes	Yes
[52]	No	Yes	W_O	Yes	Yes
[53]	No	Yes	W_O	Yes	Yes

Abbreviations: J_1: secrecy, J_2: integrity, J_3: authentication, J_4: confidentiality, J_5: authorization, J_S: proposed scheme, Yes: complying to the properties of security, No: non-complying to the properties of security, R: mutual, W_O: one way.

are unsuitable for sensitive e-healthcare applications. Adversaries can invade and obtain unlawful resources due to flaws in the structure of existing methods. Furthermore, traditional systems have high processing and transmission costs, depriving tiny, intelligent nodes of valuable resources. E-healthcare apps require a sophisticated, authenticated vital agreement method to safeguard the network against unauthorized misuse.

The proposed scheme's advantages and disadvantages are as follows:

(1) Resilient against repeated attacks.

(2) Protected against "man-in-the-middle" (MITM) attacks.

(3) Protected from attacks that modify the data.

(4) The proposed method protects the confidentiality of the data.

(5) The proposed scheme demonstrates authorization from legally responsible different stakeholders

However, the disadvantage is when used with low-power wide area networks, the proposed scheme does not result in cost saving when using a local database. This is the possibility for the scheme's future development to turn into a more affordable low-power wide area network solution. As a result of the investigation, the proposed system is seen to be superior to the traditional schemes.

3. System Design and Adversary Prototype

3.1. System Design. The admin, gateway (G_T), patient, ward boy, and doctor relationships are described in the system design. Figure 2 illustrates how stakeholders could access cloud-based health records via gateway.

3.2. Central Administrator. A hospital's IT director serves as the central administrator and registers the facility in the cloud. The administrator securely interacts with the cloud by way of the gateway using the public key of that cloud. The cloud calculates the hospital's master key and returns it to

TABLE 2: Evaluation of related study.

Framework	E-healthcare	Security risks	Awareness	Challenges	Impacts
[54]	R	ii	G	F	E
[55]	R	i	G	E	E
[56]	R	vii	F	E	G
[57]	P	ii	G	F	G
[58]	R	vi	F	E	F
[59]	I	iii	G	E	E
[60]	C	v	E	E	G
[61]	M	iv	G	F	E
[62]	M	i	G	E	F

Abbreviations: R: RFID, P: pacemaker, I: IP, M: embed medical devices, C: embed cardiac defibrillators, i: identification issues, ii: radio frequency attack, iii: intercepting attack, iv: device duplication issue, v: electromagnetic interruption, vi: illegal remote surveillance, E: high, F: medium, G: lower.
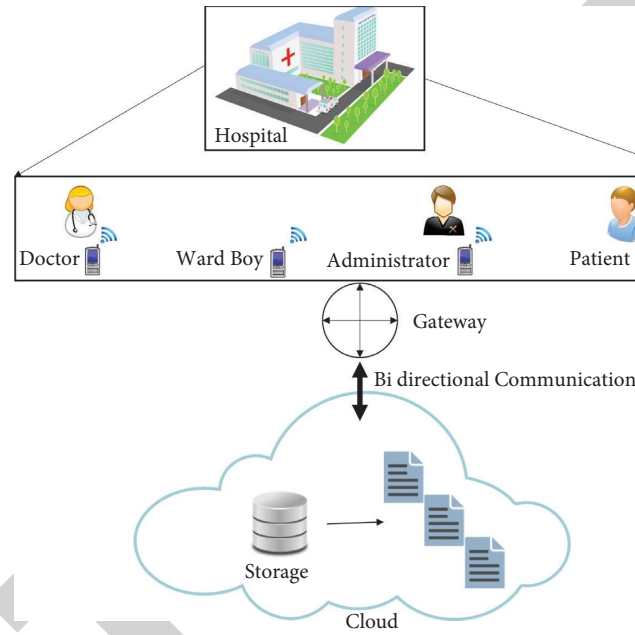
FIGURE 2: E-healthcare system based on the cloud for secure sharing with different entities.

the administrator upon registration. Subsequently, the administrator uses KCF (key conclusion function) to produce numerous subkeys from the master key. The administrator also registers the doctor, ward boy, and patient's devices offline and assigns subkeys to them.

3.3. Doctor.
A doctor is responsible for treating the patients who have been assigned to him. The patient's data should only be accessible to the doctor who is related to the patient. Comparing the cloud-based patient ID here to the patient ID provided by the doctor is how it is accomplished, and if a match is found, the request is approved; otherwise, the request is denied. Access to editing/writing of a patients' medical record is available to the doctor based on his treatment. The doctor uses encryption, hashing, and subkeying to save all of the patient's data on the cloud. The information can only be read by those who have been granted access. The doctor provides the administrator with two unique identifiers, ID_{UG} and ID_{UH}. They are kept in the cloud by the administrator. Cloud generates a one-of-a-kind DR_{ID} number. To connect securely with the gateway, a doctor uses the confidential subkey issued by the admin.

3.4. Patient.
The hospitalized person for a diagnosis or examination is referred to as the patient. A specific ward boy and doctor are allocated to care for patient in the hospital based on patient diagnosis. The patient additionally gives the administrator two unique identifiers supplied by the government (ID_{UG}) and by the hospital (ID_{UH}). The administrator gets both identifiers and keeps them in the cloud. The administrator receives the unique patient identifier (PT_{ID}) from the cloud. A patient must have the administrator's secret subkey to safely communicate with the gateway. We can assume that the devices of the stakeholders are resource constrained.

3.5. Ward Boy.
After the doctor has left, the patient is cared for by a ward boy. Ward boy gets the data from the cloud through a gateway using his subkey K_S. The ward boy supplies her government-issued unique identity number (ID_{UG}) and the hospital's address when registering offline (ID_{UH}). The ward boy receives the administrator's secret subkey K_S and the cloud's NS_{ID} (ward boy ID). A ward boy can only see the information about the patients who have been allocated to him through the healthcare department.

The patient's EHR does not allow a ward boy to update or write information in it; thus, he is only able to view the data rather than alter it. Ward boy uses his private subkey issued by the administrator to securely interact with gateway, and it is considered that user's devices have limited resources.

3.6. Relative. Relative is role which patient from its dashboard can assign. Patient can give specific access rights of its e-healthcare records stored on cloud environment from its dashboard. If patients are in critical condition or cannot share their health records from a third-party cloud, relative can access their health records and consult a doctor.

3.7. Gateway. When a user wants to access the cloud, they can do so through a gateway. The gateway is resource-unrestricted in the present system paradigm, centered on the hospital's applications. The gateway creates a secure interface for the doctor, ward boy, and patient that too view their cloud records after receiving their complete security credentials from the administration. The gateway obtains the master key K_M, and subkeys K_S and HP_{ID} through the administrator during offline registration. The doctor, ward boy, and patient communicate with the gateway via distinct subkeys while the gateway and cloud communicate via the master key (K_M); the gateway uses encryption to protect their communications.

3.8. Adversary Prototype. To disrupt normal network and service operations, hostile nodes are employed. This new protocol's ability to withstand hostile operations was evaluated using the Dolev–Yao adversary model (DY model). Messages transmitted between a gateway, a user, and the cloud can be intercepted by an adversary, according to the DY model. A hacker can intercept and replay user authentication messages en route to the cloud to illegally access cloud services. Intercepted communications may reveal secret credentials that the adversary can use for attacks like man-in-the-middle or known key emulation. An opponent can also launch a DoS attack by repeatedly bombarding the cloud.

4. Proposed Scheme for Secure Access

For any hospital, keeping track of patients and their data is a major concern for the staff; therefore, they have to multitask at all times. The solution we provided was for medical records to be stored in the cloud, saving hospital staff time and effort by eliminating the need to manually enter data. We consider the following assumptions in order to run the proposed methodology:

    (i) Although the user's device has limited computational and storage resources, the cloud and gateway are well-known entities with enormous compute and storage capacities.

    (ii) The user device, cloud, and the gateway can perform cryptographic functions.

    (iii) Data can only be accessed by a user who has been authenticated in the cloud.

The key conclusion function (KCF) is a cryptographic function for generating more than one secret keys using a master key (KCF). KCF can be used to stretch keys to make them longer or to get the keys in the desired format. The key derivation function, in this case, KCF, serves as good example of pseudorandom function. In this case, the derived key is D_K and the key conclusion function is KCF. The suggested approach comprises four stages: registration for hospital, data retrieval, data storage, and offline registration.

4.1. Registration for Hospital. The notations used all over the paper are listed in Table 3. Figure 3 depicts the administration's cloud registration process with the hospital via gateway. The nonce (N_{CC1}) is generated by administrator, who concatenates these values $RR_N \| RP_N \| HP_{ID} \| N_{CC1}$ for formation of ε. The message ε is encrypted with CK_{PU} forming v, and the resulting message (M_{01}) is forwarded to the gateway. The message v is received by the gateway, by which the nonce is generated N_{CC2}, which gets encrypted using CK_{PU} to create ι. The encrypted message gets in series with v to α, and subsequently, the generated message M_{02} is delivered to the cloud. The message received gets decrypted with CK_{PR} for computing β, and then, N_{CC2} is generated. The cloud checks the nonce N_{CC2}'s freshness. The procedure is continued if N_{CC2} is fresh; otherwise, it is aborted. To compute F, the cloud uses CK_{PR} to decrypt the message v. Cloud also checks the nonce N_{CC1} for freshness; if it is, the process is continued; otherwise, it is cancelled. The cloud checks $RR_N \| RP_N \| HP_{ID}$ and aborts the procedure if they are not found to be true. The master key K_M and nonce N_{CC3} and N_{CC4} are now generated by the cloud. To compute G, all values are concatenated $HP_{ID} \| K_M \| N_{CC3}$. By concatenating and hashing, H_S ($RR_N \| RP_N \| HP_{ID} \| N_{CC1}$), cloud additionally computes KY_{AT} at this point. A key, KY_{AT}, is used to encrypt the computed value G. $KY_{GT} = H_S$ is obtained by using hash function (N_{CC2}). To construct L, the gained key, KY_{GT}, is used for the encryption of the nonce N_{CC4}. The messages L and K are concatenated and then stored in M. The gateway receives the message M. The KY_{GT} gets calculated by the gateway using the hash of N_{CC2}. KY_{GT} is used to decrypt the message L, resulting in R. The purity of N_{CC4} is tested, and if it is found to be true, the process is restarted; otherwise, it is stopped. The value K is sent to the admin as a message M_{04} by the gateway. To construct KY_{AT}, administrator evaluates the hash value of $DR_{ID} \| PT_{ID} \| D_{RQ} \| N_{CC4}$. To create Y, the received message K is decrypted with KY_{AT}. The purity of nonce N_{CC3} is tested at this point, and if it is determined to be pure, the procedure is resumed. Finally, the administrator can successfully obtain the master key (K_M). This master key is a secure communication between the gateway and the cloud through the secret key.

4.2. Offline Registration. Registration method of offline device registration is shown in Figure 4. The administrator keeps track of each stakeholder's unique identifying

Table 3: Description of annotations.

Annotations	Description
$\|$	Concatenation operation
KY_{GT}	Temporary key of gateway
KY_{AT}	Temporary key of administration
H_S	Hash
S_{NO}	Serial number
D_S	Data to be stored
RR_N	Reference receipt number
RP_N	Payment receipt number
D_{YP}	Decryption
E_{YP}	Encryption
N_{CC}	Nonce
CK_{PU}	Public key of cloud
CK_{PR}	Private key of cloud
ID_{UG}	Unique ID issued by government
ID_{UH}	Unique ID issued by hospital
K_M	Master key
K_S	Subkey
D_{FK}	Key derivation function
D_{RQ}	Requested data
DR_{ID}	Doctor ID
PT_{ID}	Patient ID
HP_{ID}	Hospital ID
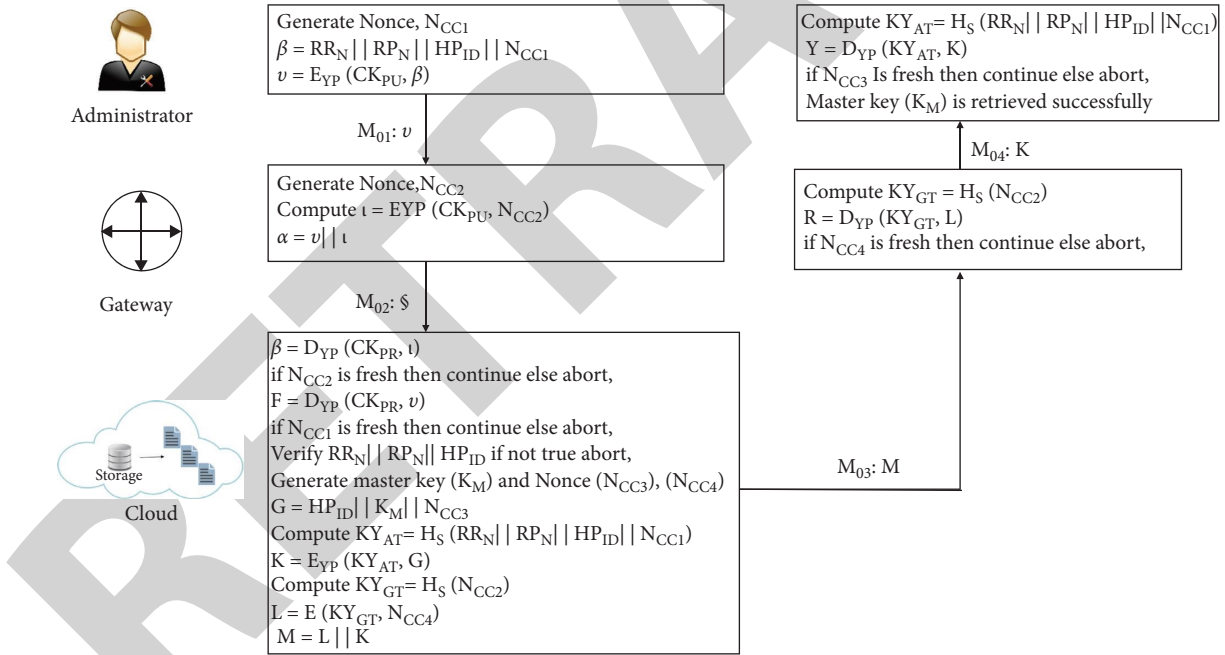
Figure 3: Registration for hospital over cloud.

information, such as ID_{UG} and ID_{UH}. As identifying data, the gateway provides the administrator with the MAC address and serial number S_{NO}. After recording the data, the administrator uploads it to the cloud, which produces the K_M and unique identifiers for the doctor $(DR_{I\,D})$, patient $(PT_{I\,D})$, and ward boy $(NS_{I\,D})$ and sends them to the admin. KCF is used by the administrator to generate several subkeys (K_S) for the secure communication between the gateway and other organizations. Users (patient, doctor, etc.) receive identification details and unique secret subkeys from the administrator, while the gateway gets the K_M, K_S,

$DR_{I\,D}$, $PT_{I\,D}$, $NS_{I\,D}$, and $HP_{I\,D}$. The unique IDs assist the administrator in ensuring the privacy of the patient's records during the offline registration phase. Only those ward boys and doctors treating that patient are given access to the patient by the administrator. The suggested approach allows the administrator to pick which stakeholders can access the cloud-based information. Table 4 shows the administrator's default settings, which include granting access and storage privileges to doctors treating the patient. On the other hand, other stakeholders have merely been provided access to the information.

$\text{ID}_{\text{UG}}, \text{ID}_{\text{UH}}$

$\text{NS}_{\text{ID}}, \text{K}_\text{S}$

Ward Boy

Administrator

$\text{DR}_{\text{ID}}, \text{K}_\text{S}$

$\text{ID}_{\text{UG}}, \text{ID}_{\text{UH}}$

Doctor

$\text{ID}_{\text{UG}}, \text{ID}_{\text{UH}}$

$\text{PT}_{\text{ID}}, \text{K}_\text{S}$

Patient

MAC, S_{NO}

$K_M, HP_{ID}, K_S, PT_{ID}, DR_{ID}, NS_{ID},$

Gateway

ID_{UG} : Unique ID issued by Government
ID_{UH} : Unique ID issued by Hospital
NS_{ID} : Ward Boy
K_S : Sub Key
PT_{ID} : Patient ID
DR_{ID} : Doctor ID
S_{NO} : Serial Number
K_M : Master Key
HP_{ID}: Hospital ID
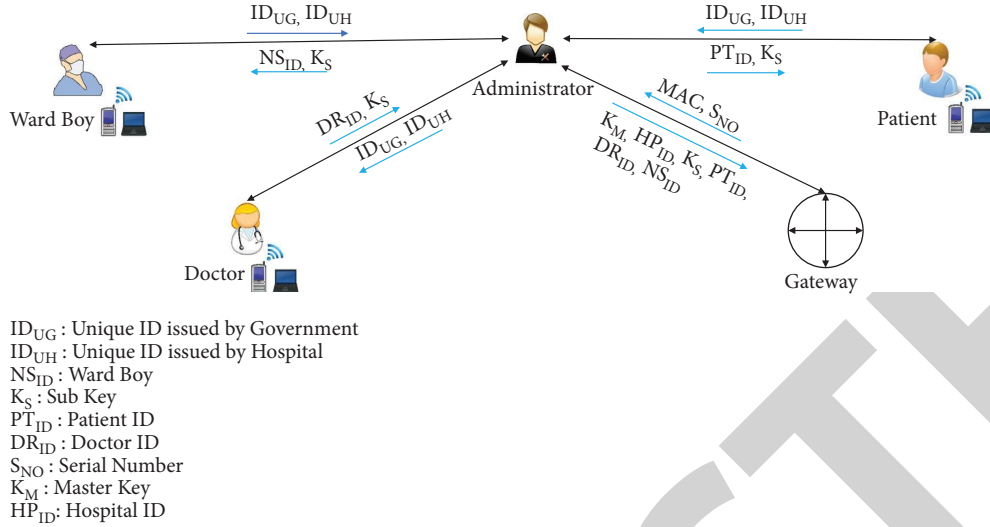
FIGURE 4: Device registration (offline).

4.3. Data Retrieval Phase. It is shown in Figure 5 that the doctor visits the gateway to express interest in communicating with a cloud-based service. The nonce N_{CC1} is generated by the doctor's device, and it is concatenated with DR_{ID} and PT_{ID} to compute γ. For computing Γ, the resulting message γ is encrypted (K_S, γ). $H_S(DR_{ID})$ has been calculated and saved in O using the hash function. Γ is combined with the message O for producing A. The gateway receives the message $S1$ from the user, which contains the value A. The gateway retrieves DR_{ID} from the database, calculates its hash value, and saves it in Z. To determine the right subkey for decryption, the gateway compares Z with O $(Z == O)$ using the subkey K_S, to form δ, and Γ is decrypted. The gateway examines the nonce N_{CC1} for freshness; if it is fresh, then the procedure is restarted; otherwise, it is aborted. The nonce N_{CC2} is generated by the gateway and is concatenated with some other values $HP_{ID}\|DR_{ID}\|PT_{ID}\|N_{\text{CC2}}$ and form ζ. Subsequently, M_K is used for the encryption of ζ in order to form η. The gateway has now sent message $S2$ to the cloud. Cloud decrypts the message η using M_K after receiving it to give ϑ. The procedure is continued if N_{CC2} is fresh; otherwise, it is aborted. The variables HP_{ID}, PT_{ID}, and DR_{ID} are checked, and if they are determined to be false, the procedure is aborted. It is checked to see whether PT_{ID} belongs to DR_{ID}, and if it does, the procedure is continued. Nonce (N_{CC3}) and requested data (D_R) are generated after successful verification and are concatenated with other values $HP_{ID}\|PT_{ID}\|DR_{ID}\|D_R\|N_{\text{CC3}}$ to form Θ. To generate κ, the Θ calculated is encrypted with M_K. The gateway receives message $S3$ from the cloud. When gateway receives a message, it decrypts it using $D_{\text{YP}}(K_M, \kappa)$ to form μ. When N_{CC3} is checked, if it is found to be fresh, the operation is continued; otherwise, it is paused. The gateway now verifies HP_{ID} and generates N_{CC4} as a nonce. Subsequently, $\varnothing$ is calculated by concatenating all of the values $DR_{ID}\|PT_{ID}\|D_{RQ}\|N_{\text{CC4}}$, and then, $\varnothing$ is encrypted using K_S to form $\varnothing$. The doctor receives the message $S4$ from the gateway and decrypts the message $\varnothing$ using K_S and computes

TABLE 4: Access rights distribution.

Device	Read	Write
Patient	Yes	No
Doctor	Yes	Yes
Ward boy	Yes	No

ρ. Only if nonce N_{CC4} is still alive, the operation is continued further. The doctor can successfully retrieve the requested data, D_{RQ}, after verifying the freshness.

4.4. Data Storage Phase. It is shown in Figure 5 that the doctor visits the gateway to engage in communicating with a cloud-based service. The nonce N_{CC1} is generated by the doctor's device, and it is concatenated with DR_{ID} and PT_{ID} to compute γ. For computing Γ, the message resulting γ is then encrypted (K_S, γ). $H_S(DR_{ID})$ has been calculated and saved in O using the hash function. Γ is merged with the message O to produce A. The gateway receives message $S1$ from the user, which contains the value A. The gateway extracts DR_{ID} through the database for calculating and storing its hash value in Z. To determine the right subkey for decryption, the gateway compares Z with O $(Z == O)$. The gateway examines the nonce N_{CC1} for freshness; if it is fresh, the procedure is restarted; otherwise, it is aborted. The nonce N_{CC2} is generated by the gateway and is concatenated with some other values $HP_{ID}\|DR_{ID}\|PT_{ID}\|N_{\text{CC2}}$ and form ζ. Subsequently, M_K has been used to encrypt ζ for formation of η. The message $S2$ is now sent to the cloud by the gateway. Cloud decrypts the message η using M_K after receiving it to give ϑ. The procedure is continued if N_{CC2} is fresh; otherwise, it is aborted. The variables HP_{ID}, PT_{ID}, and DR_{ID} are checked, and if they are determined to be false, the procedure is aborted. It is checked to see whether PT_{ID} belongs to DR_{ID}, and if it does, the procedure is continued. Nonce (N_{CC3}) and requested
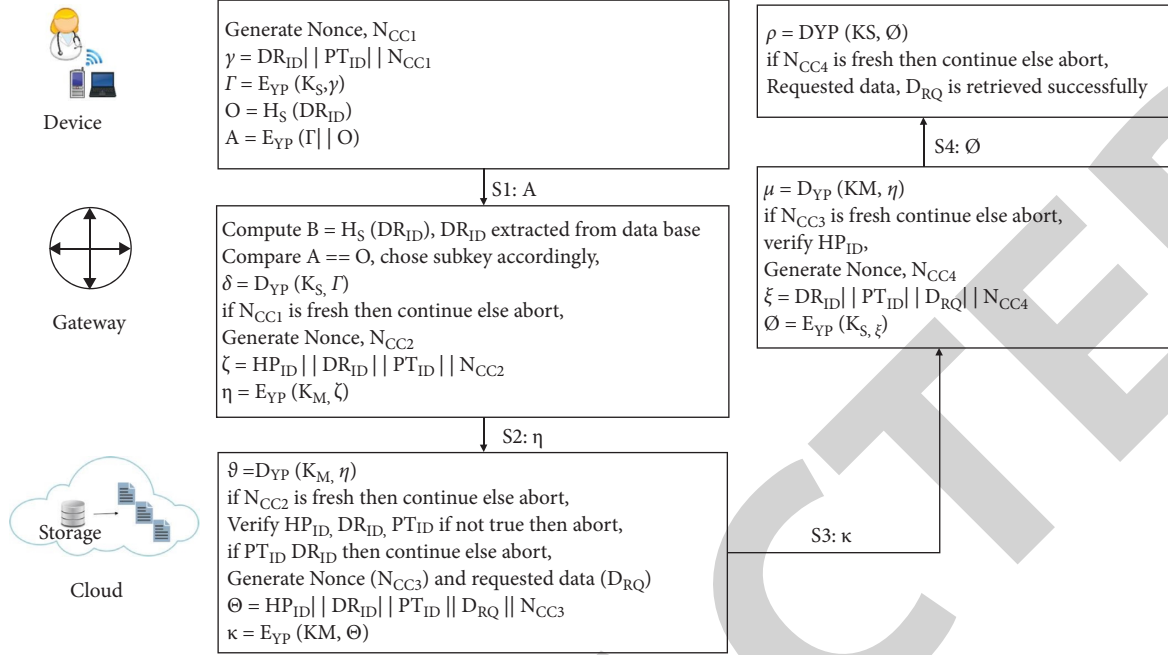
Figure 5: Retrieval phase for data.

data (D_R) are generated after successful verification and are concatenated with other values $HP_{ID}\|PT_{ID}\|DR_{ID}\|D_R\|N_{CC3}$ to form Θ. To generate κ, the Θ calculated is encrypted with M_K. The gateway receives message S3 from the cloud. When gateway receives a message, it decrypts it using $D_{YP}(K_M, \kappa)$ to form μ. When N_{CC3} is checked, if it is found to be fresh, the operation is continued; otherwise, it is paused. The gateway now verifies HP_{ID} and generates N_{CC4} as a nonce. Subsequently, $\varnothing$ is calculated by concatenating all of the values $DR_{ID}\|PT_{ID}\|D_{RQ}\|N_{CC4}$, and then, $\varnothing$ is encrypted using K_S to form $\varnothing$. The doctor receives the message S4 from the gateway after that. The doctor decrypts message $\varnothing$ using K_S and computes ρ. Only if nonce N_{CC4} is still alive, the operation is continued further. After verifying the freshness, the doctor can successfully retrieve the requested data, D_{RQ}.

5. Experimental Setup

We have developed a cloud-based web application hosted on third-party AWS cloud infrastructure. Web application is developed using PHP version 7.3. The main aim of this web application is to store and protect the e-healthcare records of the patient over a third-party cloud. The web application uses a dynamic access control scheme for sharing and uploading e-healthcare data over the cloud. Our web application service was hosted and tested on Elastic Compute Cloud (EC2) instances provided on Amazon Web Services (AWS). On the cloud platform, the virtual environment is provided by an instance of Amazon's Elastic Compute Cloud (EC2). We used EC2 instances to conduct experiments in our experimental evaluation, as shown in Table 5.

6. Experimental Design

We ran three sets of experiments to assess the web application's performance. When we conduct an experiment, we use a preallocated Amazon EC2 instance to run an artificial workload to measure the throughput (requests per second) and average web application response time. For artificial workload generation, we have used [63] for web application service performance measurement. It is possible to create an artificial user session to mimic the process of searching for an existing patient and adding a new patient record to the system. It is possible to simultaneously add a new patient to the system while conducting a patient search that returns a significant number of results from the database. The details of all the three experiments are described in Table 6.

6.1. Experiment 1: An EC2 Medium Instance with a Predetermined Allocation.
As part of experiment 1, EC2 instances assigned to the web and database tiers are shown in the following Figures 6–8 which show throughput with average response time and CPU use for the EC2 instances. The figure shows that by the 15th minute of the study, the throughput has stopped rising linearly, significantly increasing the application's response time. An obvious constraint has been discovered in the web server tiers that CPU use approaches 100%, indicating a bottleneck. After 29th minutes, the web server tier instance has stopped responding, and we cannot receive throughput and response time measures after that point in time. Howsoever, we can still monitor both instances of CPU consumption with the Amazon CloudWatch. This experiment achieved a maximum throughput of 934 requests per second.

Table 5: Resource allocation and cost.

Type of instances	vCPU (core)	RAM	Storage	Cost in (USD/hour)
m3.medium	1	3.12	4	0.071
m3.large	2	7.35	32	0.137
m3.xlarge	4	16	600	0.476

Table 6: Details on the experiment.

Number	Experiment	Description
1	Using the EC2 medium instance for a fixed amount of storage	Preallocated EC2 instances of type m3.medium and m3.2xlarge are preallocated in the web service layer and the database layer, respectively
2	Statically allocating resources with an Amazon EC2 large instance	We have preallocated two EC2 instances, one for web services and the other for databases, both of which are of type m3.2xlarge
3	EC2 xlarge instance for static allocation	Preallocated EC2 instances of type m3.2xlarge for the database and one of type m3.2xlarge for the web service tier
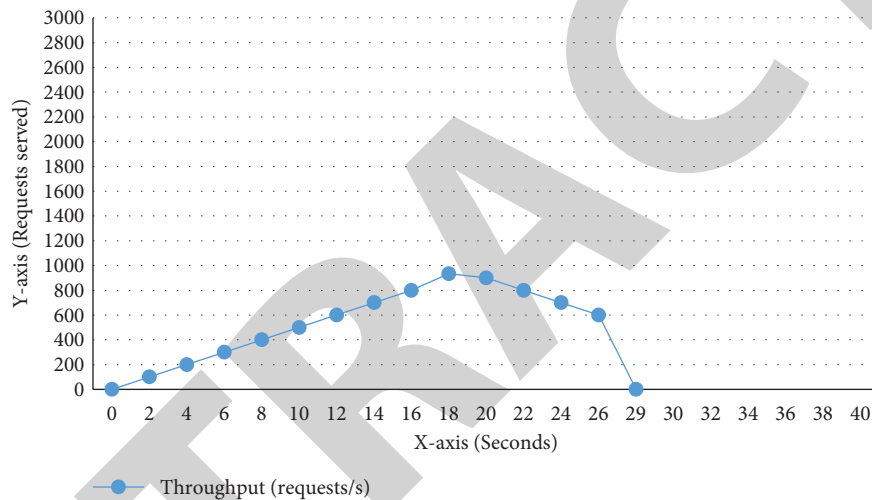
Figure 6: Throughput (requests per second).
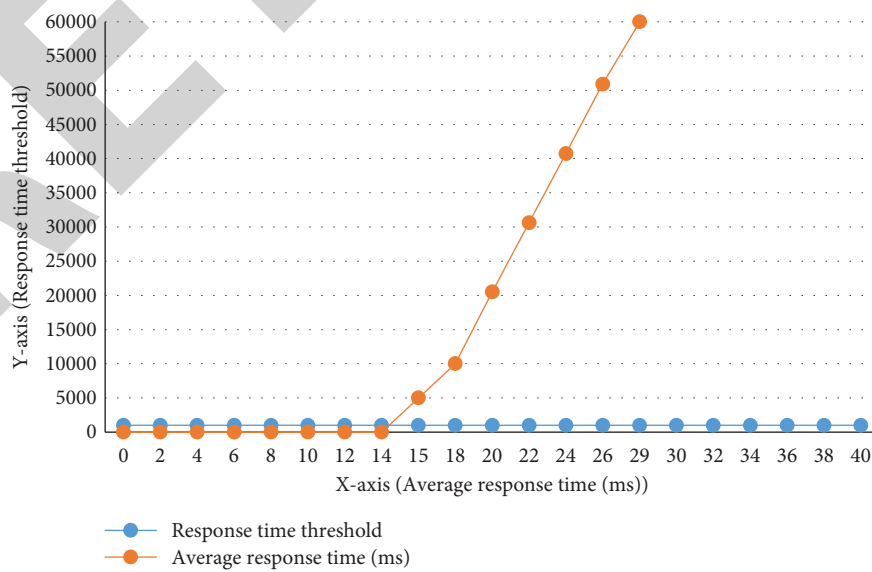
Figure 7: Avg. response time.
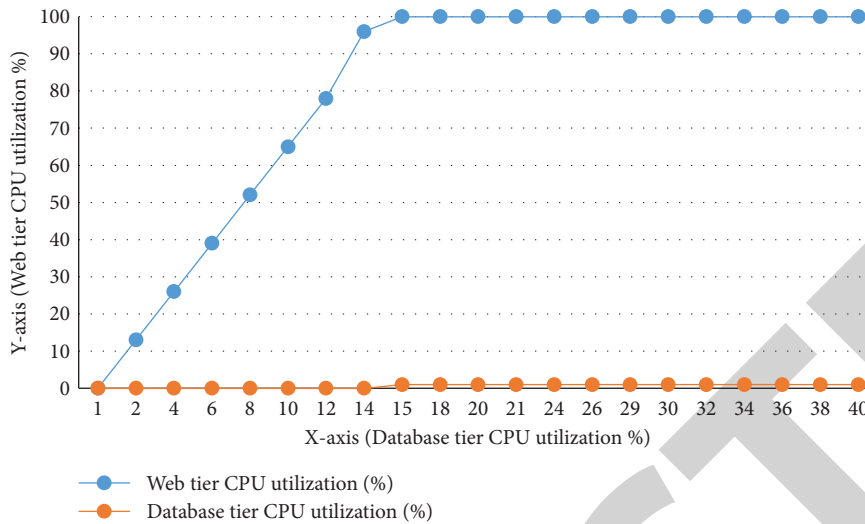
FIGURE 8: Web server and database tier instances' CPU utilization.

6.2. Experiment 2: An EC2 Large Instance with a Pre-determined Allocation. Figures 9–11 illustrate throughput, average response time, and CPU consumption for EC2 instances assigned to the web and database tiers in Experiment 2. The throughput stops increasing linearly by the 21th minute of the experiment, and there was no significant increase in response time throughout this experiment. As of this writing, the typical response time remains under around 45 milliseconds. There has been no discernible increase in CPU use in the web server with database tier replicas. This experiment gained a maximum throughput of 1051 requests per/second. Because bandwidth became a big issue at the 21th minute of the experiment and the web server instance is using 245 MB/seconds and 486 MB/second on an average for network input and output, ideally throughput should have continued to rise throughout the experiment.

6.3. Experiment 3: An EC2 xlarge Instance with a Pre-determined Allocation. Data from Experiment 3's EC2 instances are shown in Figure 12 for the throughput, average response time, and CPU utilization of the web and database tiers of EC2 instances, respectively, as depicted in Figures 13 and 14. In this experiment, there is no hint of a significant rise in response time at the point where throughput stops rising linearly at 21th minute. As of this writing, the typical response time is around 45 milliseconds. A lack of CPU saturation has been found within web and database tiers of the application. In this experiment, we achieved a maximum throughput of 1150 requests/per second. Because we see the identical bandwidth constraint in this experiment as we did in Experiment 2, the results are very similar. The conclusion is clear: increasing the web tier instance's resources does nothing to alleviate bandwidth constraints.

6.4. Comparing Cloud-Based Web Application with Block-chain Platform. EHR and EMR interoperability and security issues have been addressed by blockchain technology, which

has seen tremendous growth in the healthcare industry. Before blockchain can realize its full potential and be used in medical care, it must overcome various barriers. In this section, in terms of latency, throughput, and execution time, we test and compare the performance of the blockchain platform and web application hosted on the third-party cloud platform AWS using a secure scheme implemented for patient health records in the cloud environment. To evaluate our cloud-based web application with the Ethereum blockchain platform, we compared execution time, throughput, and latency of paper [64]. Paper [64] deployed Ethereum smart contracts using IoT devices. The purpose of paper [64] is to evaluate, store, and access transactions. We compare our patient's records, which are stored and accessed from a third-party cloud with paper [64]. The main focuses for comparing cloud-based web application with Ethereum blockchain platform are as follows:

(1) Ethereum and cloud-based web application comparison analysis.

(2) Performance matrices based on latency, throughput, and execution time are used to analyze the number of user transactions.

JMeter tools are being used in an experiment to protect patient e-health records hosted on third-party cloud providers' AWS. We then examined paper [64] Ethereum performance. We assessed latency, execution time, and throughput by different loads, such as the number of queries fired by the user on Amazon AWS cloud, for the performance evolution of cloud and blockchain technologies. The system configuration is shown in Table 7 for comparing blockchain platforms and third-party cloud platform AWS. For comparing the cloud-based web application system, we store and access e-health records of a patient analyzed with store and access transaction of Ethereum blockchain platform using IoT devices. Figures 15–19 show analysis Figures 20 of latency, execution time, and throughput of store and access transaction of third-party cloud AWS and Ethereum with IoT devices. We can see as per Figures 15 and
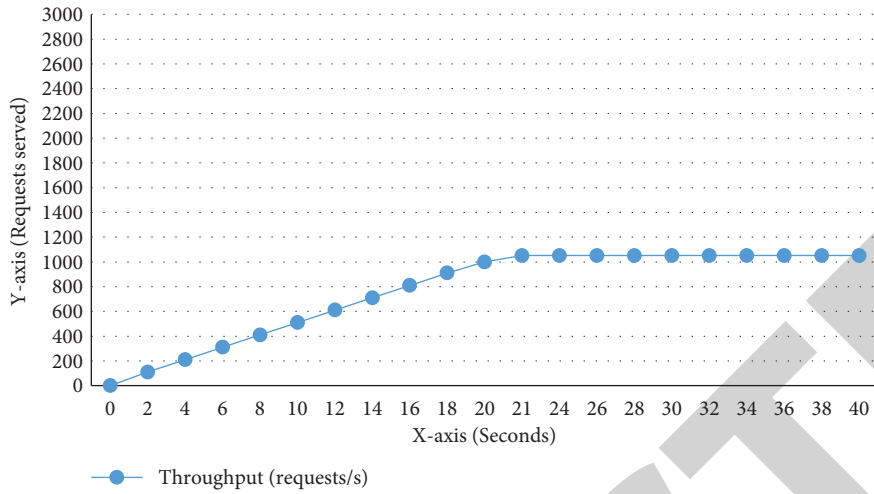
Figure 9: Throughput (requests per second).
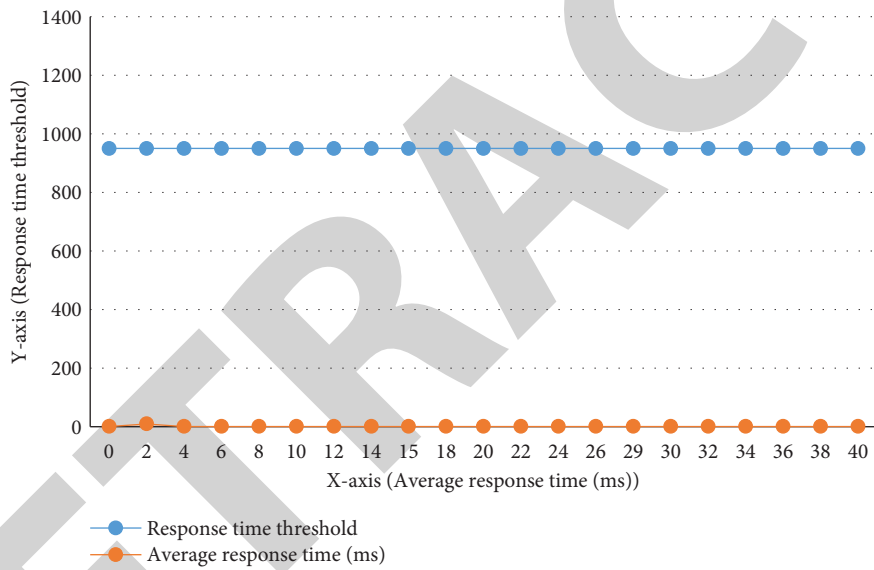
Figure 10: Avg. response time.
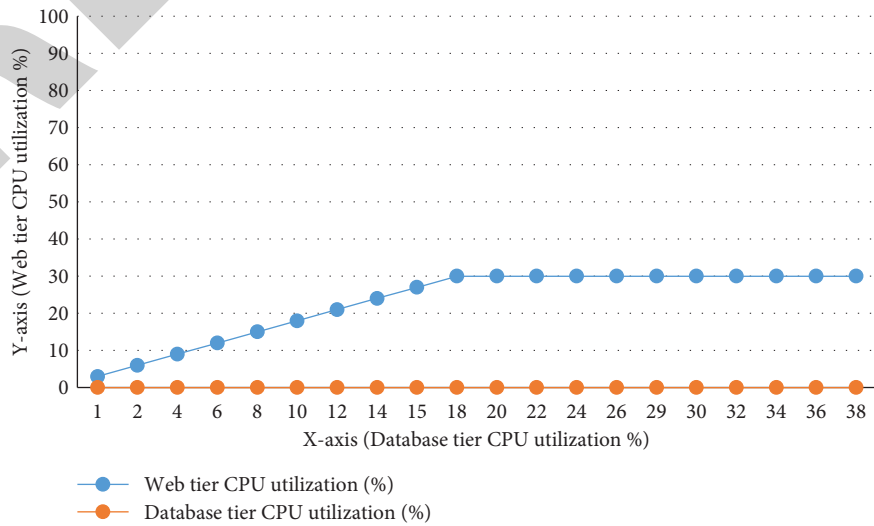
Figure 11: Web server and database tier instances' CPU utilization.
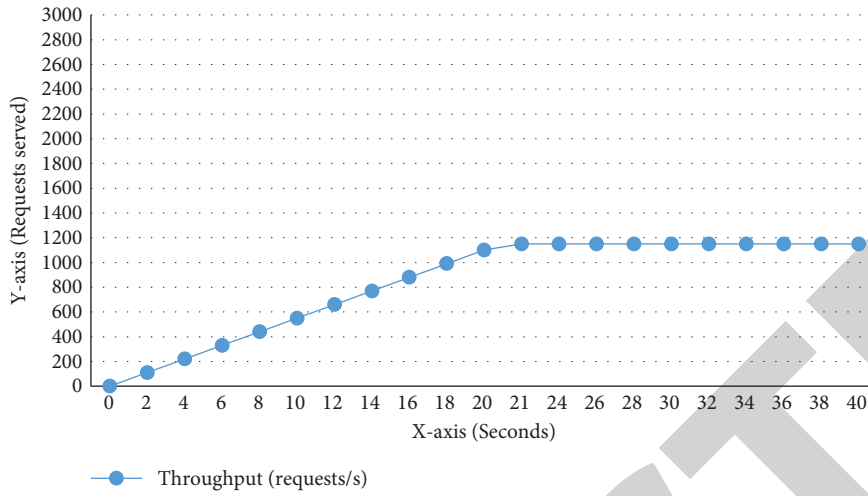
Figure 12: Throughput (requests per second).
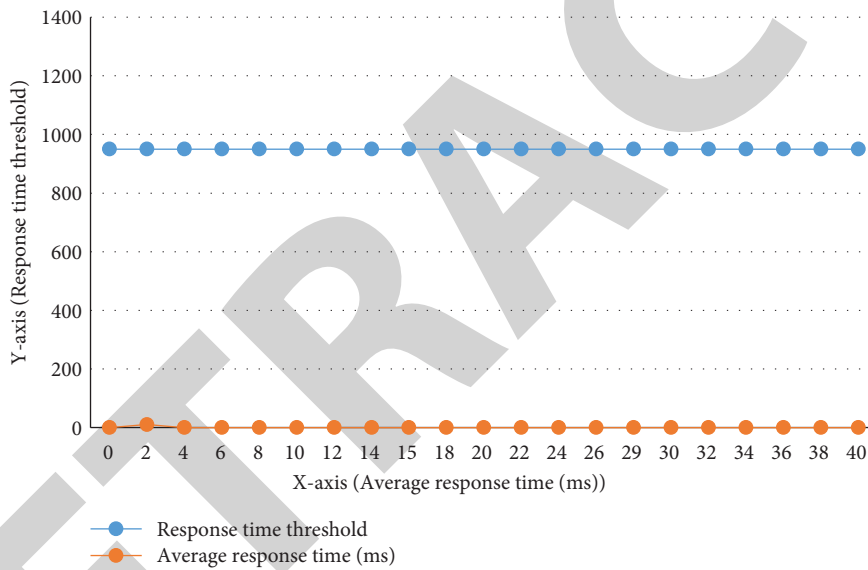
Figure 13: Avg. response time.
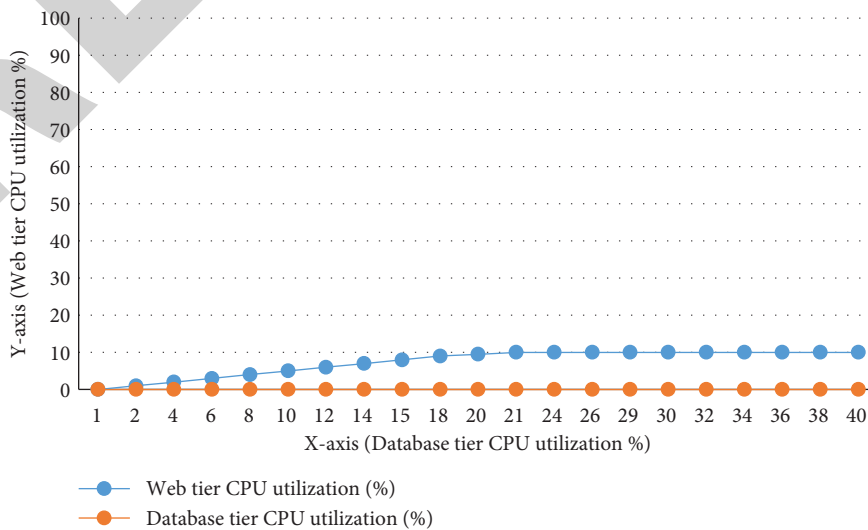
Figure 14: Web server and database tier instances' CPU utilization.

Table 7: System configuration.

Parameter	AWS	Ethereum 1.8.3
RAM	16 GB	16 GB
OS	Ubuntu Linux	Raspberry PI 3
Virtual CPU	4	8
Apache2	2.4.38	2.4.52
Storage	600 GB	1 TB
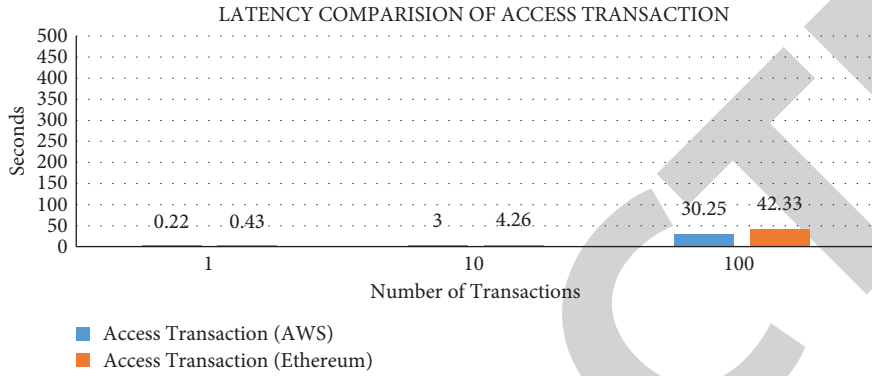
Figure 15: Latency comparison of access transaction.
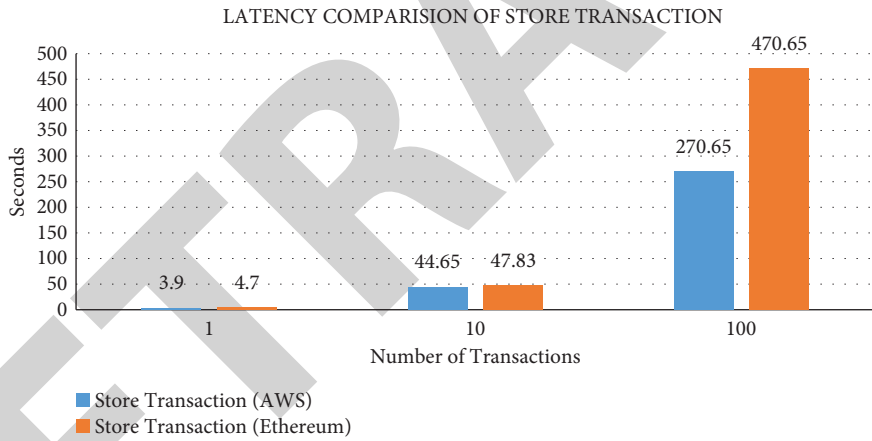
Figure 16: Latency comparison of store transaction.
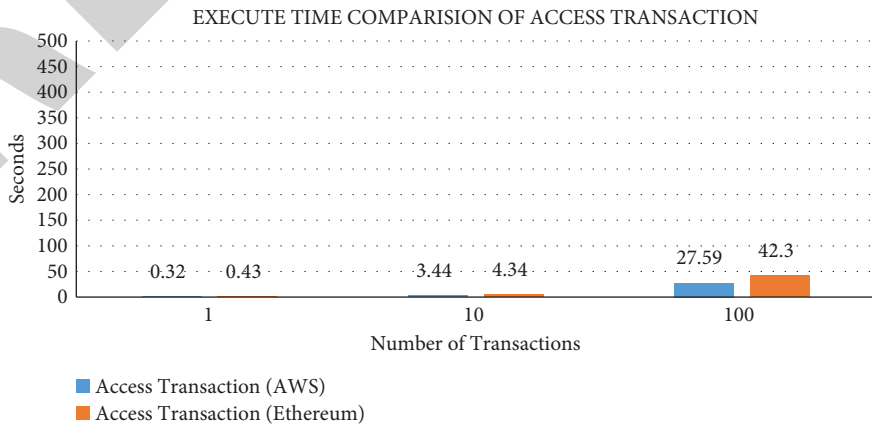
Figure 17: Execute time comparison of access transaction.

EXECUTE TIME COMPARISION OF STORE TRANSACTION
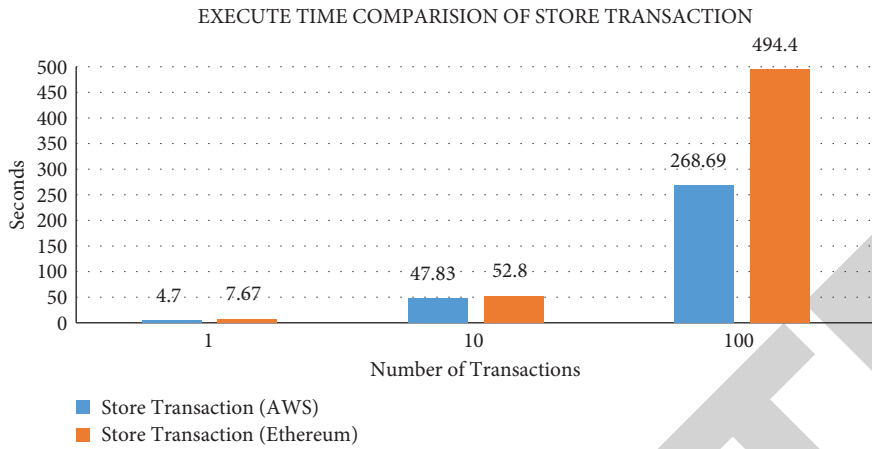
Figure 18: Execute time comparison of store transaction.
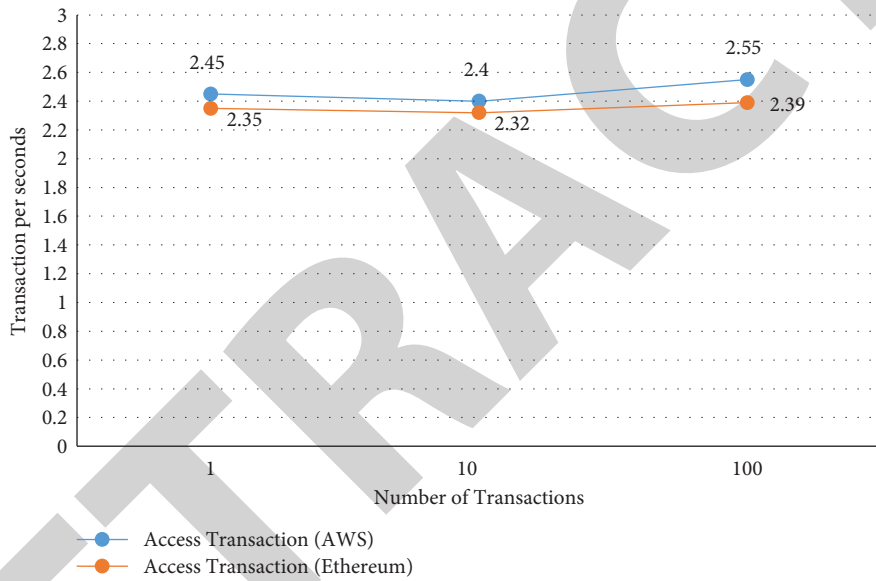
Figure 19: Throughput comparison of access transaction.
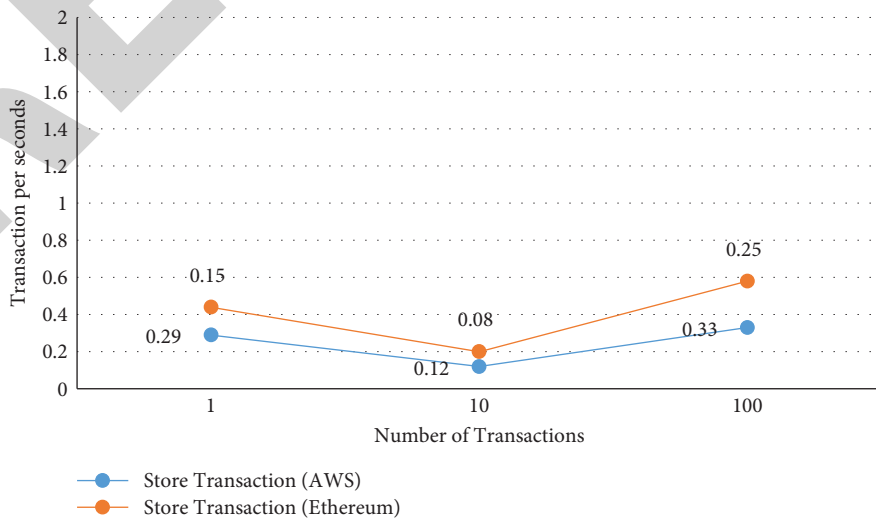
Figure 20: Throughput comparison of store transaction.

16 the latency of access and store transaction comparison. As the dataset rises in size, so does the latency of both platforms, which can be seen in the graph. As transaction gets increases latency performance of the third-party cloud-based platform is less compared to the Ethereum blockchain platform, increasing the number of transactions prompts us to look into the impact on transaction execution time variances that might exist. On both systems, as the number of transactions in the dataset grows, Figures 17 and 18 execution times are longer. Finally, the number of completed transactions per second, starting with the first deployment time, is used to determine throughput. Here, in Figures 19 and 20, it is shown that when the number of transactions is varied, the throughput remains relatively constant as the number of transactions increases; the throughput of access transactions is greater than that of store transactions in both the platforms. As transaction increases, the throughput performance of a third-party cloud-based platform is greater compared to the Ethereum blockchain platform. Secured and reliable dynamic access control scheme of patient e-healthcare records implemented in third-party cloud-based e-health application performs better in cloud environment compared to Ethereum blockchain platform using IoT devices. Finally, based on JMeter tool results, we can deduce that a cloud-based web application hosted on AWS for secure sharing of patient health records over a third-party cloud platform has the edge over the Ethereum blockchain platform.

7. Conclusions

To demonstrate the performance of our web application service, we used Amazon EC2 instances of various sizes to simulate growing workloads. We believe this study will assist web application service providers in utilizing proper cloud resources to provide response time guarantees while minimizing operational costs. Patient medical records can be easily accessed from anywhere with cloud-based e-healthcare services. Since cloud service providers offer cost-effective solutions, cloud-based e-health care has become possible. Despite the many benefits, the cloud storage and retrieval framework are particularly sensitive to wireless channels. Patient data can only be stored and accessed by those who have been granted permission (such as the patient, ward boys, doctors, and close family members). The hospital's responsibility for keeping records of patients is reduced, and access to storage of health records is improved. The proposed scheme's reliability against numerous significant attacks such as message alteration, MITMA, and replay, among others, was revealed by security analysis. The method has enormous potential for cloud-based solutions. The proposed secure access methodology for storing and accessing patient's e-health records over third-party clouds is compared to the performance evaluation of two natural traffic flow, store, and access transactions proposed using the Ethereum blockchain platform with IoT device. It can be seen that, despite having a standard system configuration, the cloud platform performs significantly better than Ethereum in relation to execution time, throughput, and

latency. Eventually, in future work, we want to test newer versions of Hyperledger Fabric with a cloud-based solution and look into different scenarios like how having numerous functions in the network affects the network's overall performance of both platforms. Furthermore, we want to compare the performance of cloud-based e-health solutions for the patient with different public blockchain technology for a higher number of transactions. Other analyses related to data security and privacy are on our agenda for the near future, particularly in the context of external access to e-healthcare data transferred through various networks over the cloud.

Data Availability

No data were used to support the findings of the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] N. Deepa, Q.-V. Pham, D. C. Nguyen et al., "A survey on blockchain for big data: approaches,opportunities,and future directions," *Future Generation Computer Systems*, vol. 131, pp. 209–226, 2022.

[2] H. K. Thakkar, P. K. Sahoo, and P. Mohanty, "Dofm: domain feature miner for robust extractive summarization," *Information Processing & Management*, vol. 58, no. 3, Article ID 102474, 2021.

[3] R. Kumar, M. Swarnkar, G. Singal, and N. Kumar, "Iot network traffic classification using machine learning algorithms: an experimental analysis," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 989–1008, 2021.

[4] C. S. Kruse, M. Mileski, A. G. Vijaykumar, S. V. Viswanathan, U. Suskandla, and Y. Chidambaram, "Impact of electronic health records on long-term care facilities: systematic review," *JMIR medical informatics*, vol. 5, no. 3, p. e7958, 2017.

[5] A. Mubashar, K. Asghar, A. R. Javed et al., "Storage and proximity management for centralized personal health records using an ipfs-based optimization algorithm," *Journal of Circuits, Systems, and Computers*, vol. 31, no. 01, Article ID 2250010, 2022.

[6] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5g," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2399–2406, 2017.

[7] Y. Zhang, X. Ma, J. Zhang, M. S. Hossain, G. Muhammad, and S. U. Amin, "EdGe intelligence in the cognitive internet of things: Improving Sensitivity and Interactivity," *IEEE Network*, vol. 33, no. 3, pp. 58–64, 2019.

[8] A. Ghoneim, G. Muhammad, S. U. Amin, and B. Gupta, "Medical image forgery detection for smart healthcare," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 33–37, 2018.

[9] M. Chen, J. Yang, L. Hu, M. S. Hossain, and G. Muhammad, "Urban healthcare big data system based on crowdsourced and cloud-based air quality indicators," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 14–20, 2018.

[10] G. S. Gaba, G. Kumar, H. Monga, T.-H. Kim, M. Liyanage, and P. Kumar, "Robust and lightweight key exchange (lke) protocol for industry 4.0," *IEEE Access*, vol. 8, Article ID 132808, 2020.

[11] W. Min, B.-K. Bao, C. Xu, and M. S. Hossain, "Cross-platform multi-modal topic modeling for personalized inter-platform recommendation," *IEEE Transactions on Multimedia*, vol. 17, no. 10, pp. 1787–1801, 2015.

[12] S. Hiremath, G. Yang, and K. Mankodiya, "Wearable Internet of Things: Concept, Architectural Components and Promises for Person-Centered Healthcare," in *Proceedings of the 2014 4th international conference on wireless mobile communication and healthcare-transforming healthcare through innovations in mobile and wireless technologies (MOBIHEALTH)*, pp. 304–307, IEEE, Athens, Greece, November 2014.

[13] M. Masud, M. S. Hossain, and A. Alamri, "Data interoperability and multimedia content management in e-health systems," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 6, pp. 1015–1023, 2012.

[14] A. Abbas and S. U. Khan, "A review on the state-of-the-art privacy-preserving approaches in the e-health clouds," *IEEE journal of Biomedical and health informatics*, vol. 18, no. 4, pp. 1431–1441, 2014.

[15] M. AlOtaibi, A. T. Lo'ai, and Y. Jararweh, "Integrated Sensors System Based on Iot and mobile Cloud Computing," in *Proceedings of the 2016 IEEE/ACS 13th international conference of computer systems and applications (AICCSA)*, pp. 1–5, IEEE, Agadir, Morocco, November 2016.

[16] M. F. Alhamid, M. Rawashdeh, H. Al Osman, M. S. Hossain, and A. El Saddik, "Towards context-sensitive collaborative media recommender system," *Multimedia Tools and Applications*, vol. 74, no. 24, pp. 11399–11428, 2015.

[17] P. Kumar and G. S. Gaba, "Biometric-based robust access control model for industrial internet of things applications," *IoT Security: Advances in Authentication*, pp. 133–142, Wiley Telecom, Hoboken, 2020.

[18] S. Chenthara, K. Ahmed, H. Wang, and F. Whittaker, "Security and privacy-preserving challenges of e-health solutions in cloud computing," *IEEE Access*, vol. 7, Article ID 74361, 2019.

[19] R. Charanya and M. Aramudhan, "Survey on access control issues in cloud computing, in: 2016 international conference on emerging trends in engineering," in *Proceedings of the Technology and Science (ICETETS)*, pp. 1–4, IEEE, Pudukkottai, India, February 2016.

[20] W. Li, K. Xue, Y. Xue, and J. Hong, "Tmacs: a robust and verifiable threshold multi-authority access control system in public cloud storage," *IEEE Transactions on parallel and distributed systems*, vol. 27, no. 5, pp. 1484–1496, 2015.

[21] K. Xue, Y. Xue, J. Hong et al., "Raac: robust and auditable access control with multiple attribute authorities for public cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 4, pp. 953–967, 2017.

[22] P.-W. Chi and C. L. Lei, "Audit-free cloud storage via deniable attribute-based encryption," *IEEE Transactions on Cloud Computing*, vol. 6, no. 2, pp. 414–427, 2015.

[23] W. Li, B. M. Liu, D. Liu et al., "Unified fine-grained access control for personal health records in cloud computing," *IEEE journal of biomedical and health informatics*, vol. 23, no. 3, pp. 1278–1289, 2018.

[24] C. Zhang, L. Zhu, C. Xu, and R. Lu, "Ppdp: an efficient and privacy-preserving disease prediction scheme in cloud-based e-healthcare system," *Future Generation Computer Systems*, vol. 79, pp. 16–25, 2018.

[25] C. Huang, K. Yan, S. Wei, G. Zhang, and D. H. Lee, "Efficient Anonymous Attribute-Based Encryption with Access Policy Hidden for Cloud Computing," in *Proceedings of the 2017 international conference on progress in informatics and computing (PIC)*, pp. 266–270, IEEE, Nanjing, China, December 2017.

[26] J. Wei, W. Liu, and X. Hu, "Secure data sharing in cloud computing using revocable-storage identity-based encryption," *IEEE Transactions on Cloud Computing*, vol. 6, no. 4, pp. 1136–1148, 2016.

[27] H. Cui, R. H. Deng, and Y. Li, "Attribute-based cloud storage with secure provenance over encrypted data," *Future Generation Computer Systems*, vol. 79, pp. 461–472, 2018.

[28] L.-C. Huang, H.-C. Chu, C.-Y. Lien, C.-H. Hsiao, and T. Kao, "Privacy preservation and information security protection for patients' portable electronic health records," *Computers in Biology and Medicine*, vol. 39, no. 9, pp. 743–750, 2009.

[29] Al. Hamid, S. M. M. Rahman, M. S. Hossain, A. Almogren, and A. Alamri, "A security model for preserving the privacy of medical big data in a healthcare cloud using a fog computing facility with pairing-based cryptography," *IEEE Access*, vol. 5, Article ID 22313, 2017.

[30] M. Masud, G. S. Gaba, K. Choudhary, R. Alroobaea, and M. S. Hossain, "A robust and lightweight secure access scheme for cloud based e-healthcare services," *Peer-to-peer Networking and Applications*, vol. 14, no. 5, pp. 3043–3057, 2021.

[31] M. S. Hossain and G. Muhammad, "Cloud-based collaborative media service framework for healthcare," *international journal of distributed sensor networks*, vol. 2014, no. 4, Article ID 858712, 2014.

[32] F. Whittaker, "Security and privacy-preserving challenges of e-health solutions in cloud computing," *IEEE Access*, vol. 7, Article ID 74361, 2019.

[33] D. R. Matos, M. L. Pardal, P. Adao, A. R. Silva, and M. Correia, "Securing electronic health records in the cloud," in *proceedings of the 1st workshop on privacy by design in distributed systems*, pp. 1–6, New York, NY, USA, April 2018.

[34] F. Rezaeibagha and Y. Mu, "Distributed clinical data sharing via dynamic access-control policy transformation," *International Journal of Medical Informatics*, vol. 89, pp. 25–31, 2016.

[35] M. Marwan, A. Kartit, and H. Ouahmane, "A cloud based solution for collaborative and secure sharing of medical data," *International Journal of Enterprise Information Systems (IJEIS)*, vol. 14, no. 3, pp. 128–145, 2018.

[36] H. Zhang, J. Yu, C. Tian, P. Zhao, G. Xu, and J. Lin, "Cloud storage for electronic health records based on secret sharing with verifiable reconstruction outsourcing," *IEEE Access*, vol. 6, Article ID 40713, 2018.

[37] H. Elmogazy and O. Bamasak, "Towards Healthcare Data Security in Cloud Computing," in *Proceedings of the 8th international conference for internet technology and secured transactions (ICITST-2013)*, pp. 363–368, IEEE, London, UK, December 2013.

[38] M. K. Hasan, S. Islam, I. Memon et al., "A novel resource oriented dma framework for internet of medical things devices in 5g network," *IEEE Transactions on Industrial Informatics*, p. 1, 2022.

[39] M. K. Hasan, M. Shafiq, S. Islam et al., "Lightweight cryptographic algorithms for guessing attack protection in complex internet of things applications," *Complexity*, vol. 202113 pages, Article ID 5540296, 2021.

[40] S. Amanlou, M. K. Hasan, and K. A. A. Bakar, "Lightweight and secure authentication scheme for iot network based on

Research Article

Pigeon Inspired Optimization with Encryption Based Secure Medical Image Management System

B. T. Geetha,[1] Prakash Mohan,[2] A. V. R. Mayuri,[3] T. Jackulin,[4] J. L. Aldo Stalin,[5] and Varagantham Anitha [6]

[1]*Department of ECE, Saveetha School of Engineering, SIMATS, Saveetha University, Chennai, India*
[2]*Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, India*
[3]*SCSE, VIT Bhopal University, Bhopal, India*
[4]*Department of CSE, Panimalar Engineering College, Chennai, India*
[5]*Department of Information Technology, Sona College of Technology, Salem, India*
[6]*Department of Computer Science, Institute of Technology, Hawassa University, Hawassa, Ethiopia*

Correspondence should be addressed to Varagantham Anitha; anithav@hu.edu.et

Presently, technological advancements in the healthcare sector pose a challenging problem relevant to the security and privacy of health-related applications. Medical images can be considered significant and sensitive data in the medical informatics system. In order to transmit medical images in an open medium, the design of secure encryption algorithms becomes essential. Encryption can be considered one of the effective solutions for accomplishing security. Although numerous models have existed in the literature, they could not adaptable to the rising number of medicinal images in the health sector. At the same time, the optimal key generation process acts as a vital part in defining the performance of the encryption techniques. Therefore, this article presents a Pigeon Inspired Optimization with Encryption-based Secure Medical Image Management (PIOE-SMIM) technique. The proposed PIOE-SMIM approach majorly concentrates on the development of secret share creation (SSC) and the encryption process. At the initial stage, the medical images are converted into a collection of 12 shares using the SSC approach. In addition, an elliptic curve cryptography (ECC) scheme is employed for the encryption process. In order to optimum key creation procedure in the ECC model, the PIO technique is exploited with the aim of maximizing PSNR. Finally, on the receiver side, the decryption and share reconstruction processes are performed to construct the original images. The PIOE-SMIM model displayed an enhanced PSNR of 59.37 dB in image 1. Improved PSNR of 59.53 dB is given for image 5 using the PIOE-SMIM model. For demonstrating an enhanced performance of the PIOE-SMIM method, a widespread experimental study is made and the results highlighted the supremacy of the PIOE-SMIM model over other techniques.

1. Introduction

Most healthcare providers in recent times have implemented any procedure of electronic mode of medical record system [1]. For reducing the cost of medical imaging (MI) structures by means of front-end ownership and IT maintenance cost burdens, many health practices have changed to outsource options and save MI data from the centralization databases on a third-party cloud which is not completely trustworthy and could be probably presented at remote areas. Though, if

the MI data is saved on a third-party cloud and probably distrusted cloud, security, and secrecy in relation to the safeguarding of data and data access become critical problems [2, 3]. Security and secrecy problems of the healthcare cloud emerge primarily from the inherent property of cloud computing (CC), namely heterogeneity and service composition and consolidation, multitenancy, and dynamic resources [4].

Initially, if the medical activities deploy the cloud memory services to save their MI data on a remote cloud [5],

these medical practices activities needed to encounter the actual risk that their MI data are being widely open to the cloud service provider by which unlawful accesses for invading the secrecy and confidentiality of the saved medicinal data [6]. Second, the medicinal data from various medicinal practices are frequently held in a similar memory cloud for data accessing and processing [7]. Confirming the security purpose and secrecy of patients and other private information is an ongoing research area. Hence, we definitely needed an encryption structure that delivers a very effective method for controlling data accessibility on the basis of user attributes (or privileges), relatively more than their identities. The healthcare cloud is a mission-critical cloud service system that needs the strongest security and secrecy guarantees for medicinal practices with reference to both data as well as access [8, 9]. First, for the sake of preventing information leakages from the saved data, all health data saved on the healthcare cloud must be encoded [10]. Pigeons are simple and intelligent birds that can fly long distances in search of food and then return home without becoming disoriented. This move encouraged researchers, who concluded that pigeons navigate by using the Earth's magnetic field, differences in sun altitude, and memory of a few landmarks. The success of any algorithm is guaranteed when the algorithm can explore and exploit the problem domain globally. Probing PIO on multiple applications allows us to gain a deeper understanding of the method. Pigeons are known to travel large distances in quest of food. They can fly through a forest as well as 759 International Journal of Engineering and Technology through wide space. Another intriguing feature is that the pigeons can find their way back to their home by sensing the Earth's magnetic field, the sun's height, and visual cues such as landmarks. Leading pigeons in the flock communicate with the rest of the flock and navigate by keeping a close flocking distance. A pigeon's leadership is defined by the number of times a specific bird speaks with the other pigeons in a given random population.

The researchers in [11] examined the safety of medicinal images in the IoT environment through the advanced cryptographic method with metaheuristic optimizer-based approaches. For growing the security level of encrypting as well as decrypting procedures, the optimum key would be selected via hybrid swarm optimization, viz., GSO, and PSO in elliptic curve cryptography. In [12], the dual encryption process is exploited for encrypting medicinal images. At first, Blowfish Encryption is taken into account and signcryption procedure is employed to authorize the encryption method. Next, the Opposition-based Flower Pollination (OFP) is exploited to upgrade the public and private keys. Blowfish algorithm encrypts data and expands keys. P–array and S–boxes are used at the beginning of key expansion with numerous subkeys, requiring precomputation before encryption or decryption. P–array has 18 32-bit subkeys, P1 to P18. This section converts 448-bit keys into 4168-byte subkey arrays—The second stage of the Blowfish algorithm encrypts data, starting with a 64-bit plaintext block and ending with a 64-bit ciphertext. First, identical 64-bit segments 39. Second, the exclusive or–operation (XOR) between the initial 32-bit block termed left (L) and the first

P-array. The 32-bit data from Step 2 is passed to the F function, which permutes it into a 32-bit block segment and XORs it with the right (R) 32-bit block of the main 64-bit plaintext. After XOR, L and R segments are exchanged for Blowfish iterations. P1 through P18 are utilized in reverse order for decryption. An opposition-based flower pollination algorithm, also known as an OBFPA, has been used to solve problems involving the optimization of functions and the construction of structures in engineering. The enhancement is accomplished through the utilization of two primary optimization methodologies. The local self-adaptive greedy approach promotes the population's ability to exploit its resources, while the global opposition-based learning of the elite population increases the population's genetic variation. The 18 benchmark functions and two structure engineering design issues are used to validate an elite opposition-based flower pollination method.

This study devises a Pigeon Inspired Optimization with Encryption-based Secure Medical Image Management (PIOE-SMIM) technique. The proposed PIOE-SMIM technique intends in the design of secret share creation (SSC) and the encryption process. At the initial stage, the medical images are converted into a collection of 12 shares using the SSC approach. In addition, an elliptic curve cryptography (ECC) scheme is employed for the encryption process. In order to produce optimum keys involved in the ECC model, the PIO algorithm is exploited with the aim of maximizing the peak signal-to-noise ratio (PSNR). Finally, on the receiver side, the decryption and share reconstruction processes are performed to construct the original images. For demonstrating an enhanced performance of the PIOE-SMIM model, a comprehensive simulation study is made. In summary, the study's contributions are given in the following.

(i) To introduce a new PIOE-SMIM model for a secure image management system in the healthcare sector

(ii) To employ the ECC-based encryption technique to securely transmit the medical images

(iii) To design a PIO algorithm to optimally generate the keys for enhanced secure medical image transmission.

(iv) To validate the performance of the proposed model against a benchmark medical dataset and examine the results interms of different measures.

The rest of the study is planned as follows. Section 3 presents the proposed PIOE-SMIM model and Section 4 elaborates the result analysis. Finally, the concluding remarks are drawn in Section 5.

2. Literature Survey

Sambit et al. [13] examined secure medicinal images with a pair of subkeys, first a pair of subkeys are provided through the tent and chaotic logistic maps. Adoptive GSO with PSNR and correlation coefficient fitness function has been presented for choosing the optimum public and secret keys amongst the arbitrary number.

Hasan et al. [14] introduced an effective, lightweight encryption model to design a secured image coding approach for the medical sector. The presented method applies two permutation approaches for securing medicinal images. The presented method is evaluated, analyzed, and compared with the encrypted one in security and implementation time. In [15], DNA-based cryptography, as well as dual hyperchaotic map methods, are suggested to offer higher-level security for the medicinal image. The digital medicinal image is larger in size and requires computation time. Song et al. [16] proposed a rapid and real-time cryptosystem by the means of chaotic encryption and parallel computing. The concept is based on a "permutation-substitution" structure of a chaotic encryption system. During the substitution phase, encryption over adjacent images through an expanded cipher-blockchain (CBC) model ensures desired diffusion in the entire image set. During the permutation phase, the batch image is classified into different groups. Then, multi-thread generates permutation coordinates simultaneously in every group to save time, from which more than one image is shuffled at one go.

The researchers in [17] employ the chaotic tent map and SCAN concept for medicinal image encryption. It comprises of pixel rearranging stage, diffusion operation, and bit plane decomposition. The amount of data in the images is divided into eight binary planes with differing proportions. Low and high four-bit planes undergo shuffling by diagonal and spiral SCAN patterns. In [18], an Improved Chaos Encryption (ICE) is employed for improving the secrecy using arbitrary nature. The overall energy is computed on the image and a comparison is made with adaptive thresholds for segmenting the Lorenz module employed in the chaotic technique for improving the security approach.

In [19], the authors presented a V-net convolution neural networks (CNNs) model relying on the 4D hyperchaotic method to encode medicinal images. At first, the plaintext medicinal image deal with 4D hyperchaotic sequence image, including pseudorandom sequence generation, image segmentation, and chaotic system processing. Next, the V-net CNN model is utilized for training chaotic sequences to remove the periodicity of chaotic sequences. At last, the chaotic sequence image undergoes diffusion for changing actual image pixels for realizing the encryption process. Masood et al. [20] a light-weighted cryptosystem based on Chen's chaotic, Henon chaotic map, and Brownian motion system is proposed for encrypting the healthcare images with higher security. The experiment result shows that the presented system is a lightweight technique that could accomplish the anticipated security level to encrypt secure image-based patient datasets.

Block-based image processing and hyper-image encryption techniques were used in this study to secure images. A transformation algorithm reorganized the original image into the transformed image and then encrypted the encrypted image using the Blowfish algorithm, or hyperimage encryption techniques. According to the results, the correlation between image elements has been much reduced. A lower correlation and higher entropy were also found when the number of blocks was increased by using smaller block sizes. There is not a key generator in this algorithm. In order to protect the image, a hyper-image encryption method was utilized. IJSER International Journal of Scientific and Engineering Research, Volume 7, Issue 1, January-2016 96 ISSN 2229–5518 is unable to process this volume's enormous amount of data. International Journal of Science and Engineering Research Multimedia data may not be acceptable for text-based algorithms. The degree of correlation between different aspects of an image was greatly reduced using this approach.

In Matlab 2013, Blowfish Algorithm is used to encrypt and decrypt images. In this study, a secret-key block cypher named 64 bits Blowfish was proposed for picture encryption and decryption, to improve both security and performance. This algorithm is used to generate keys with a range of sizes ranging from 32 bits to 448 bits. A Feistel network is used, which repeats a simple function sixteen times. There are no known vulnerabilities in the blowfish algorithm. It also outperforms the current industry standard in terms of speed and efficiency MATLAB is used to develop and implement the proposed algorithm. As a result, the blowfish algorithm becomes more robust when the number of rounds is increased. Blowfish is an effective standard encryption algorithm because it has no known security flaws.

3. The Proposed Model

In this article, a new PIOE-SMIM model has been developed for accomplishing a secure image transmission system. The proposed PIOE-SMIM technique mainly focuses on the design of SSC and the encryption process. At the initial stage, the medical images are converted into a collection of 12 shares using the SSC approach. In addition, the ECC scheme is employed for the encryption process. In order to optimize the key generation procedure in the ECC model, the PIO technique is exploited to maximize PSNR.

3.1. Share Creation Process. Generally, an image is comprised of overlapping shares, from which the human optical system finds private images without primitive technology. At last, the count of shares is widely distributed. The homogenous model contains sensor hubs that are similar because it is farther from dimensions namely battery power and the devices are extremely complex to construct. Hence, an image generates "N" shares for the security system. E-Secret picture distribution is a fine talent and skill that concerns itself with the security of significant color photographs by dividing the secret into numerous copies. This is done in order to ensure the integrity of the image. The core idea behind the technique is to convert the color image into a number of different unreadable format type shares. This will ensure that the information contained inside the image will remain hidden until it is once again combined with some sort of mathematical computation. Security, reconstruction precision, computation complexity, and storage requirements are the four conditions that can be used to evaluate a Secret Sharing Scheme's performance.

Encryption is the process of transforming information into a code that is only known to a select few, hence concealing the information's actual meaning. The process of encrypting and decrypting information is referred to as cryptography, which is a scientific discipline. In computers, data that has not been encrypted is referred to as plaintext, and data that has been encrypted is referred to as ciphertext. Decryption is the process of converting data that has been encrypted back into its original form. In most cases, decryption is simply the process of encryption done backward. It takes the encrypted information and decodes it so that only a user with the appropriate authorization may decrypt the data. This is because decryption requires a secret key or password. Encryption is the process of transforming a message that can be read into one that cannot be read, to prevent third parties who are not authorized from accessing the information. The process of turning an encrypted message back to its original format (which can be read by humans) is referred to as decryption. The communication in its original form is referred to as the plaintext message. At the initial stage, the medical images are converted into a collection of 12 shares using the SSC approach. The pixel values of images were determined and the RGB values get described by matrix (R_m, G_m, B_m). The size of the matrix is equal to input image size ($P * Q$). It is described by the following equation:

$$\text{Pixel} = \sum R + G + B, \tag{1}$$

whereas pixel denotes the overall R_m, G_m, and B_m pixel values. All the pixels present in the input images undergo n transformed way named shares. The RGB shares are depending upon the pixel value that exists in the RGB image. The shares for RGB are symbolized by R_s, G_s, and B_s.

$$R_s = \int_1^k \frac{\lim}{k \longrightarrow 1\text{ton}} R_{ab},$$

$$G_s = \int_1^k \frac{\lim}{k \longrightarrow 1\text{ton}} G_{ab}, \tag{2}$$

$$B_s = \int_1^k \frac{\lim}{k \longrightarrow 1\text{ton}} B_{ab},$$

whereas a and b shows the location of the matrices, (R_s, G_s and B_s) signifies the RGB share, (R_{ab}, G_{ab} and B_{ab}) represents the component of the image pixel. In previous to share formation, the elementary matrix is required to be derivative according to the number of shares to be produced that is fixed by the user. Here, the elementary matrix amount is 2 and the share amount is 4. Next, the elementary matrix is acquired by the abovementioned process and is characterized by B_{M1} and B_{M2} correspondingly. In previous to making share, the succeeding process is implemented on the matrices XR_1 and XR_2.

$$XR_1 = 128 - B_{M1},$$
$$XR_2 = B_{M2}. \tag{3}$$

The red band share is generated by XOR operation among the basic matrix and key matrix as shown in the following.

$$Rs1 = XR_1 \oplus K_M,$$
$$Rs2 = XR_2 \oplus XR_1,$$
$$Rs3 = XR_2 \oplus Rs_1, \tag{4}$$
$$Rs4 = Rs1 \oplus R.$$

In the reconstruction process, multiple shares are incorporated to generate the original images that are shown below:

$$R = Rs1 \oplus Rs2 \oplus Rs3 \oplus Rs4 \oplus Rs4 \oplus K_M,$$
$$G = Gs1 \oplus Gs2 \oplus Gs3 \oplus Gs4 \oplus Gs4 \oplus K_M, \tag{5}$$
$$B = Bs1 \oplus Bs2 \oplus Bs3 \oplus Bs4 \oplus Bs4 \oplus K_M.$$

When the share is recreated, the encryption and decryption method with the ECC approach is exploited on each color band of the recreated share. All the band images are separated into blocks previous to encryption and decryption procedures. The block is divided into the size of $4 * 4$. From the abovementioned process, the share is produced and an encryption procedure is employed.

3.2. ECC Based Encryption.

Once the shares are generated, the ECC scheme is employed for the encryption process. ECC is a public key cryptographic method relying on the algebraic architecture of elliptical curves through finite fields. It includes a small key through the non-EC cryptography for offering similar safety. Figure 1 displays the working process of the ECC. At the ECC, the prime value is elected as n_p and the private key is preferred as H [21]. Next, it can be expressed in the following:

$$E = p(i)^3 + u * p(i) + v, \tag{6}$$

where u and v denote the constant value $u = v = 2$. Once the state $X = Y$ is satisfied, the optimum points are selected for the ECC. Next, X and Y aredetermined as

$$X = \text{mod}(E, n_p),$$
$$Y = \text{mod}(p(j)^2, n_p). \tag{7}$$

Now, $p(i, j)$ denotes the point of the elliptic curve. n_p describes a prime number. The doubling technique is exploited for defining the X and Y values. The optimum point $P_e(k, l)$ and P_f indicates public key as given below:

$$P_f = H * P_e. \tag{8}$$

In the encryption procedure, each share comprises a block and all the blocks are comprised of encrypted sections. The total number of blocks is indicated by $b(i, j)$ whereas i and j indicates the row along with the column of the blocks. At this point, all the portions of information are delivered as an input to encrypt data. Data $D_x(i, j)$ and $D_y(i + 1, j)$ along with the point can be described as follows [22]:

$$C_1 = H * P_e,$$
$$C_2 = (D_x, D_y) + C_1. \tag{9}$$
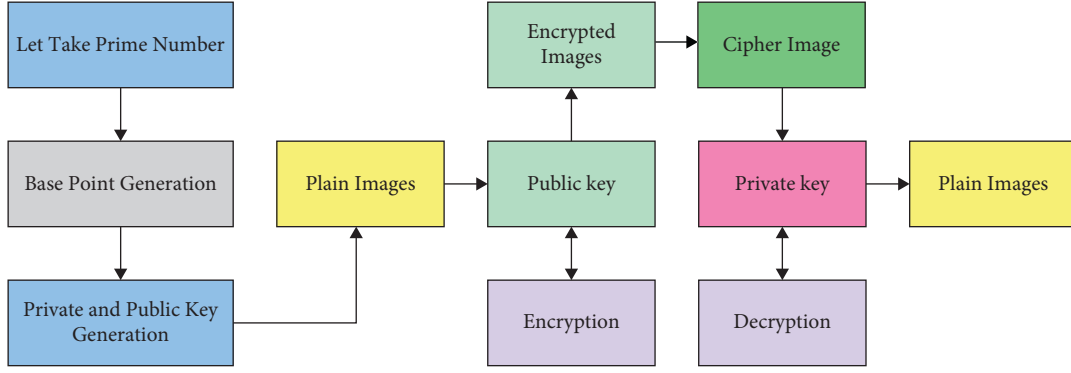
FIGURE 1: Process in ECC.

In decryption procedure, the private key (H) is employed for communication decryption whereas the point C_{11} is exploited to decrypt pixels.

$$C_{11} = H * C_1,$$
$$C_{ij} = C_2 - C_{11}. \quad (10)$$

The C_{ij} shows the final result of the decryption procedure. From the outcomes of C_{ij}, pixel values of IR and original color bands (RGB) are uniquely saved. Eventually, it can be expressed as follows:

$$F_{\text{image}} = R + G + B. \quad (11)$$

Now, the keys exploited into encrypting and decrypting data are produced through the MRFO process that is deliberated in the succeeding subsection.

3.3. Optimal Key Generation. To generate ECC keys optimally, the PIO algorithm is exploited to maximize PSNR. The POA has three operators: the compass, landmark, and map operators. During the map and compass operators, pigeons sense the geomagnetic field for the procedure the mapping to home [23]. In this recently developed technique, the map and compass operator model is presented based on the magnetic field and the sun, whereas the landmark operator model is presented based on the landmarks themselves. Map and compass operator: The birds would orient themselves by using the magnetic field of the Earth (also known as the CMT), which would also enable them to take over in the event that the environmental conditions no longer permitted optical orientation. This phenomenon is known as magnetoreception. Landmark operator: Pigeons rely on the neighboring landmarks in order to find their way as they go closer and closer to their destination. If they are familiar with the locations, they will bypass the intermediate stops and go straight to their destination. They will follow the pigeons that are familiar with the landmarks in the event that they are a considerable distance from the goal and do not know where the landmarks are located. Consider that searching space is N dimensional, and i^{th} pigeon of swarms is represented as N dimensional vectors $X_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,N})$. The velocity of pigeons

signifies the altered place of pigeons and is demonstrated as other N dimensional vectors AC. The previous visited place of ith pigeons are signified as $P_i = (P_{i,1}, P_{i,2}, \ldots, P_{i,N})$. The global optimum place of swarms is presented as $(g_1, g_2, \ldots, g_N)$. Every pigeon was flying dependent upon equations (12) and (13):

$$V_i(t+1) = V_i(t) \times e^{-Rt} + rand \times \left(X_g - X_i(t)\right), \quad (12)$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \quad (13)$$

where R implies the map and compass factors, whereas r and refers to the arbitrary value in the range of [0, 1], X_g represents the global optimum solutions, $X_i(t)$ signifies the present place of pigeons at the time t, and $V_i(t)$ indicates the existing velocity of pigeon at iterations t. During the landmark operator, each pigeon is ranked dependent on its fitness values. During every group, the amount of pigeons is upgraded by equation (14), whereas the number of pigeons were taken into account for evaluating the chosen place of centered pigeons, whereas each other pigeon changes its terminus as follows [24].

$$N_p(t+1) = \frac{N_p(t)}{2}, \quad (14)$$

where N_p signifies the number of pigeons from the existing iteration t. The place of the chosen terminus was measured by equation (1), whereas each other pigeon upgrades its place nearby equation (16).

$$X_c(t+1) = \frac{\sum X_i(t+1) \times \text{Fitnes}\left(X_i(t+1)\right)}{N_p \sum \text{Fitness}\left(X_i(t+1)\right)}, \quad (15)$$

$$X_i(t+1) = X_i(t) + rand \times \left(X_c(t+1) \times X_i(t)\right), \quad (16)$$

where X_c signifies the place of centered pigeons. Figure 2 demonstrates the flowchart of the PIO approach.

Optimization based on pigeons as seen in the images above, bio-inspired Swarm Intelligence Optimizer (SIO) is presented for the classification of satellite images [25–29]. Scientists have found that pigeons' orientation is primarily based on two operators that appear to utilizing particular rules. The pigeon's natural behavior, according to estimates, is that they travel long distances in search of food.
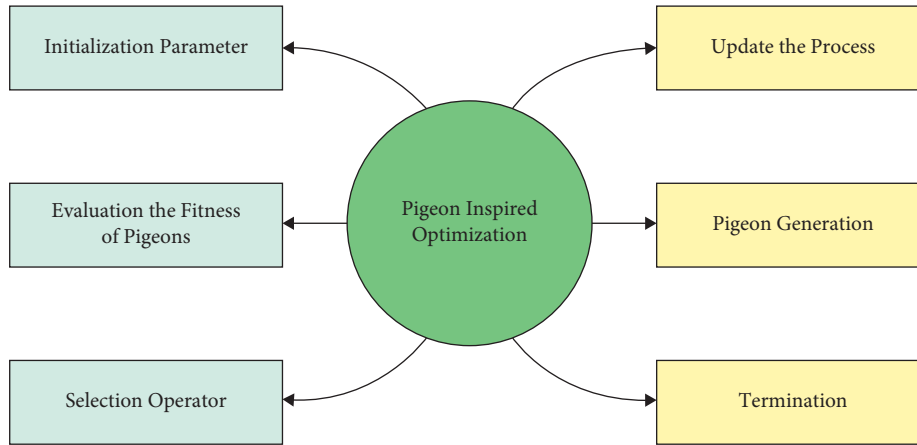
FIGURE 2: Flowchart of PIO technique.

They are equally adept at flying in dense woodland as they are at soaring through an open field. The pigeons' capacity to perceive the Earth's magnetic field, the sun's height, and visual cues like landmarks is another intriguing characteristic. Leading birds communicate and navigate with the remainder of the flock by flying side-by-side with each other. Researchers have shown that pigeons' amazing directing abilities are nearly entirely dependent on small magnetic particles in their beaks. Pigeons, in particular, have iron crystals in their beaks that give them a sense of direction. The trigeminal nerve appears to transmit messages from the nose to the brain via magnetite particles, according to research [30–33]. The pigeon's ability to detect the difference in height between the sun at the base and at the time of release has been regarded as evidence that the sun also plays a role in the pigeon's navigation. There is new evidence that pigeons can follow major highways, railroads, and waterways instead of flying straight to their destination, according to studies on the birds' behaviors. An optimum key selective method assumes that 'fitness function' is the maximal key using PSNR for scrambling and unscrambling data in the medicinal images in IOT. The procedure was generated by the model of hybrid optimized for assessment. It can be demonstrated in the subsequent condition.

$$\text{fitness} = \text{maximum } \{\text{PSNR}\}. \tag{17}$$

4. Experimental Validation

The experimental validation of the PIOE-SMIM model is tested using three medical datasets namely DR [25], dermoscopic [26], and lung [27] datasets. Figure 3 demonstrates a sample test image.

To comprehend the share conception, a group of four shares (R, G, and B band) created to perform dermoscopic images is depicted in Table 1. The three rows from the table represent the shares of the RGB band correspondingly. As the shares outperformed in the table, it can be clear that the shares are not meaningful.

For understanding the share conception, a group of four shares (R, G, and B band) created to execute the DR image is exposed in Table 2. The three rows from the table signify the shares of the RGB band correspondingly. As demonstrated in the table, it can be clear that the shares do not convey any meaning.

Table 3 provides a brief result examination of the PIOE-SMIM model under distinct test images. Figure 4 highlights a brief PSNR investigation of the PIOE-SMIM model under a distinct number of test images. The figure indicated that the PIOE-SMIM model has resulted in increased values of PSNR. For instance, with image 1, the PIOE-SMIM model has provided a PSNR of 59.37 dB. Moreover, with image 2, the PIOE-SMIM model has provided a PSNR of 59.99 dB. Similarly, with image 3, the PIOE-SMIM model has offered a PSNR of 58.82 dB. Moreover, with image 4, the PIOE-SMIM model has resulted in a PSNR of 60.28 dB. At last, with image 5, the PIOE-SMIM model has accomplished a PSNR of 59.53 dB.

Figure 5 demonstrates a detailed SSIM and CC study of the PIOE-SMIM model under a distinct number of test images. The figure designated that the PIOE-SMIM model has resulted in enlarged values of SSIM and CC. For instance, with image 1, the PIOE-SMIM model has offered SSIM and CC of 99.90 and 99.95. Also, with image 2, the PIOE-SMIM model has provided SSIM and CC of 99.99 and 99.91. Equally, with image 3, the PIOE-SMIM model has offered SSIM and CC of 99.90 and 99.97. Furthermore, with image 4, the PIOE-SMIM model has resulted in SSIM and CC of 99.94 and 99.91. Finally, with image 5, the PIOE-SMIM model has led to SSIM and CC of 99.90 and 99.96.

A detailed PSNR and MSE comparison study of the PIOE-SMIM model with other models is provided in Table 4 and Figure 6. The outcomes implied that the PIOE-SMIM model has accomplished enhanced values of PSNR and reduced values of MSE. For instance, with image 1, the PIOE-SMIM model has provided the least MSE of 0.0751 whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have accomplished increased MSE of 0.1050, 0.1254, 0.1521, and 1.4460, respectively. At the same time, with image 5, the PIOE-SMIM model has offered a reduced MSE
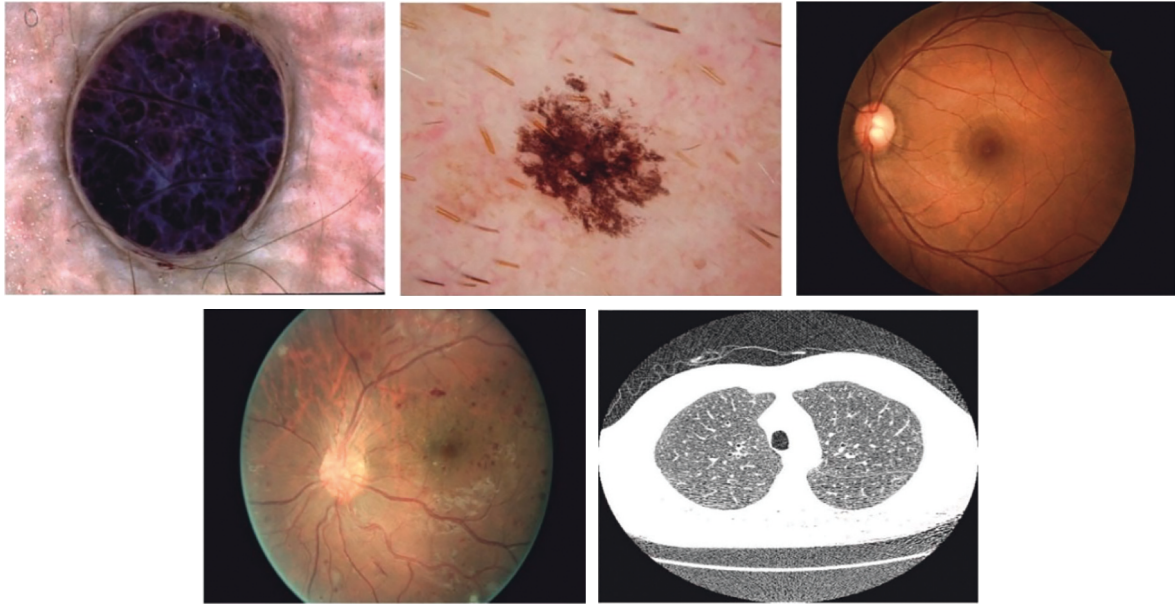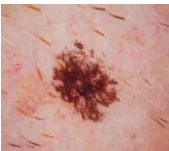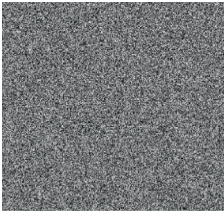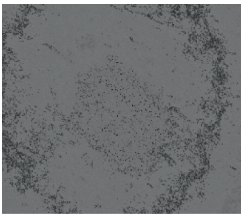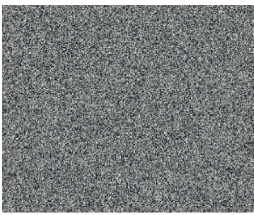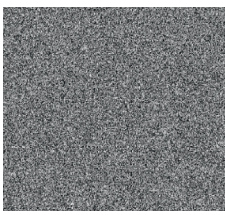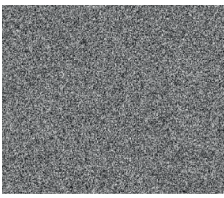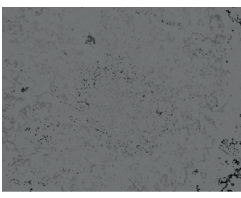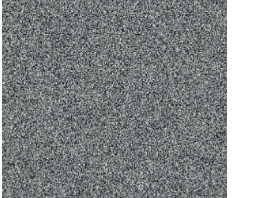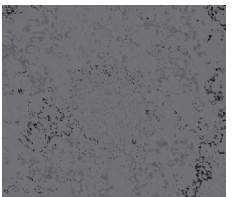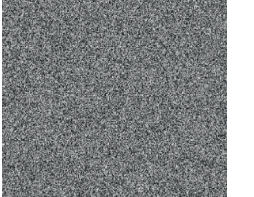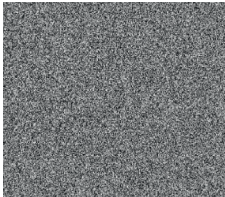
FIGURE 3: Sample medical images.

TABLE 1: Visualization of multi-share creation scheme image 1.

Test image	Share 1	Share 2	Share 3	Share 4

of 0.0725 whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have gained improved MSE of 0.0812, 0.1280, 0.1542, and 2.1652, respectively.

A thorough CC comparison study of the PIOE-SMIM model with other models is provided in Table 5 and Figure 7. The outcomes portrayed that the PIOE-SMIM model has resulted in improved values of CC.

For instance, with image 1, the PIOE-SMIM model has provided a maximum CC of 99.95 whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have accomplished minimum CC of 99.71, 98.03, 98.11,Similarly, with image and 97.68 respectively. In line with image 5, the PIOE-SMIM technique has offered superior CC of 99.96 whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models

TABLE 2: Visualization of multiple share creation scheme image 2.

Test image	Share 1	Share 2	Share 3	Share 4
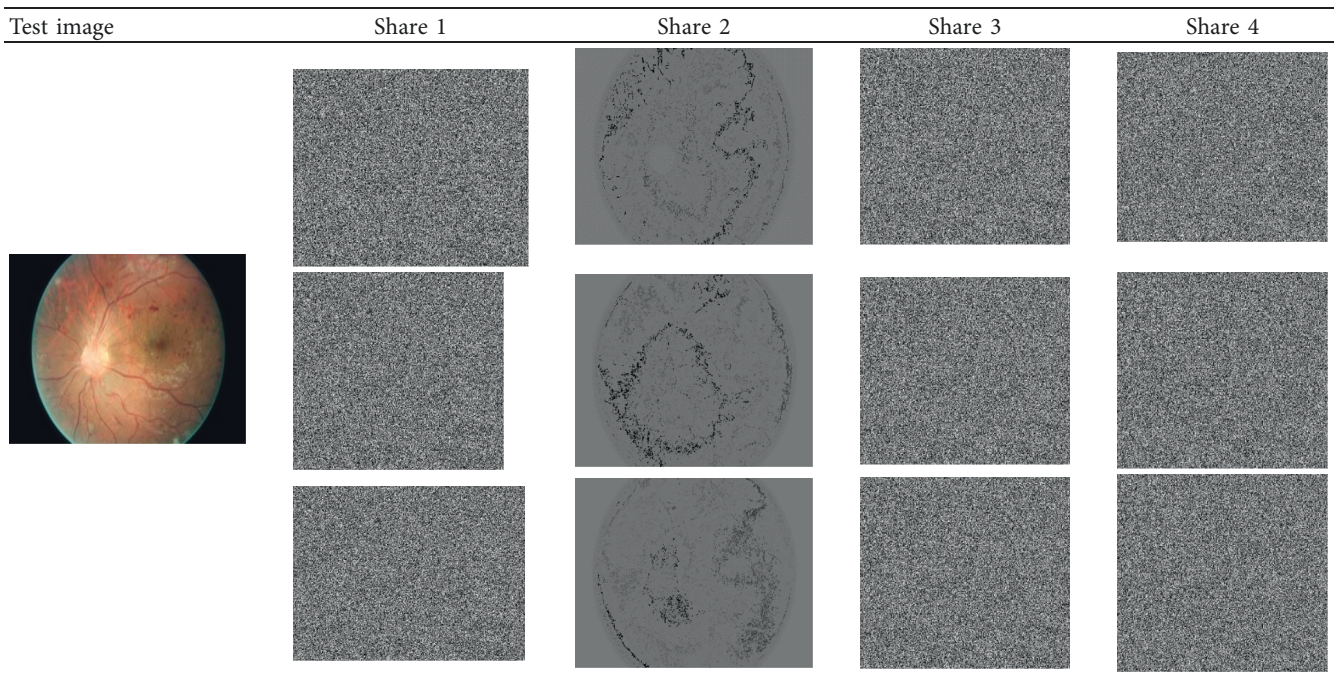
TABLE 3: Result analysis of PIOE-SMIM method with distinct measures and images.

Test images	MSE	RMSE	PSNR	SSIM	CC
Image-1	0.0751	0.2740	59.37	99.90	99.95
Image-2	0.0652	0.2553	59.99	99.99	99.91
Image-3	0.0853	0.2921	58.82	99.90	99.97
Image-4	0.0610	0.2470	60.28	99.94	99.91
Image-5	0.0725	0.2693	59.53	99.90	99.96
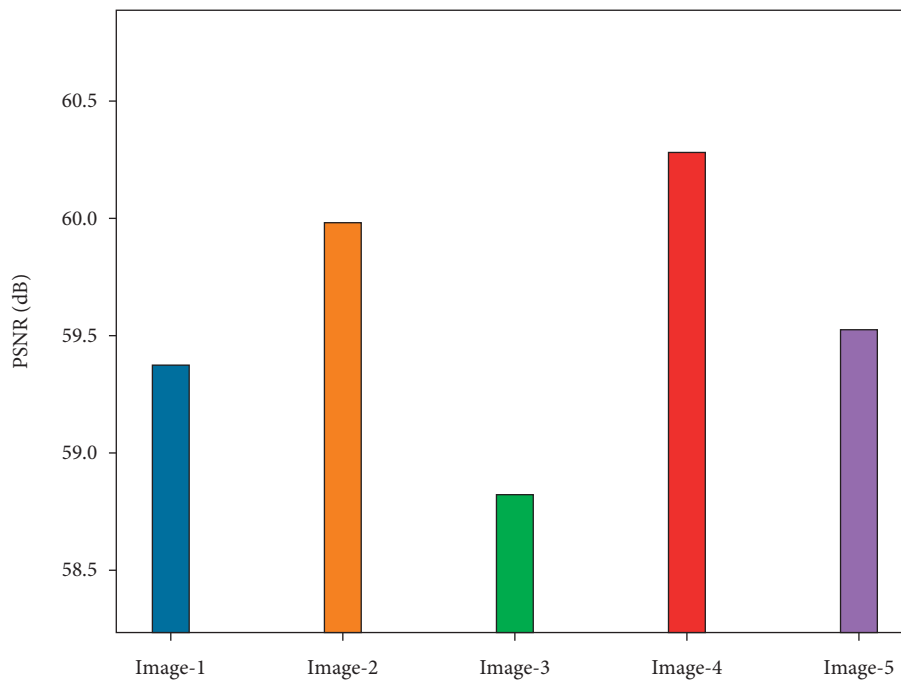
FIGURE 4: PSNR analysis of PIOE-SMIM technique with distinct images.
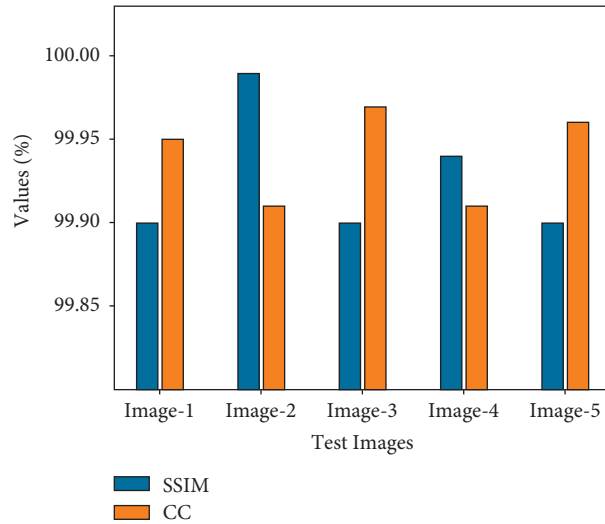
FIGURE 5: SSIM and CC analysis of PIOE-SMIM technique with distinct images.

TABLE 4: Comparative analysis of PIOE-SMIM method with current methods in terms of MSE and PSNR.

Input image	PIOE-SMIM		OSC-MI		Hybrid model		GWOE-SCO		SC-ECC	
	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR	MSE	PSNR
Image-1	0.0751	59.37	0.1050	57.92	0.1254	57.15	0.1521	56.31	1.4460	46.53
Image-2	0.0652	59.99	0.1287	57.04	0.1354	56.81	0.1479	56.43	1.9453	45.24
Image-3	0.0853	58.82	0.1054	57.90	0.1268	57.10	0.1395	56.69	2.0245	45.07
Image-4	0.0610	60.28	0.0812	59.04	0.1521	56.31	0.1952	55.23	1.7516	45.70
Image-5	0.0725	59.53	0.0812	59.04	0.1280	57.06	0.1542	56.25	2.1652	44.78
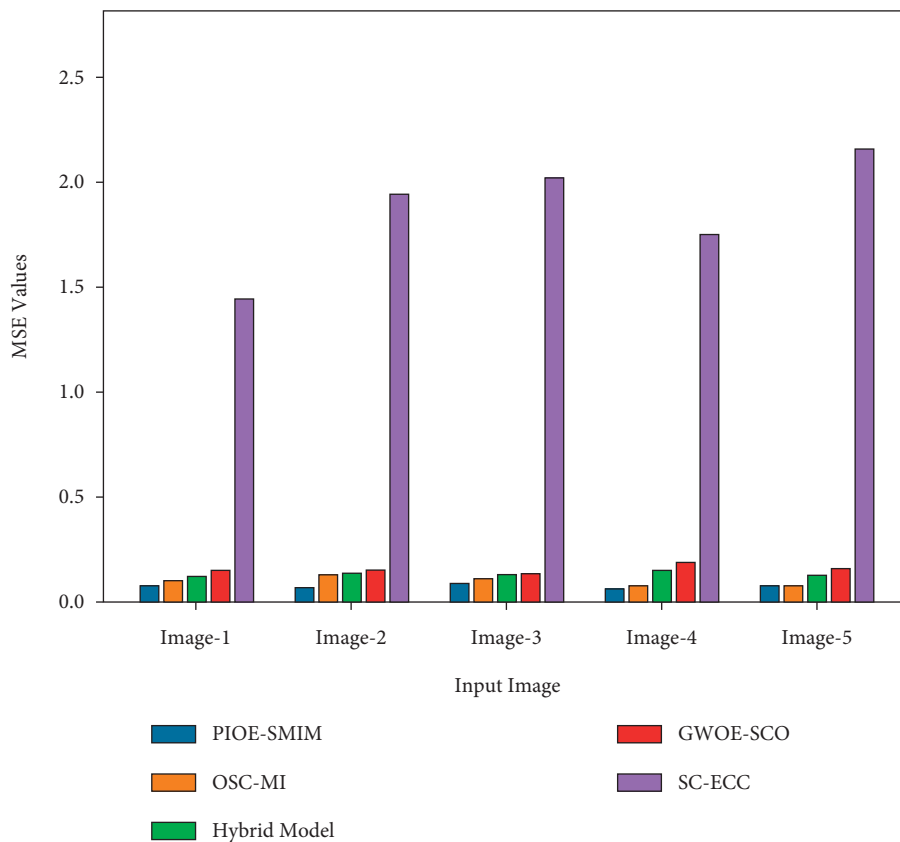
FIGURE 6: Comparative analysis of PIOE-SMIM technique with existing methods.

TABLE 5: Result analysis of PIOE-SMIM technique with recent algorithm in terms of CC.

Input image	PIOE-SMIM	OSC-MI	Hybrid model	GWOE-SCO	SC-ECC
Image-1	99.95	99.71	98.03	98.11	97.68
Image-2	99.91	99.24	98.71	98.25	98.14
Image-3	99.97	99.78	98.78	97.11	96.52
Image-4	99.91	99.61	98.71	97.32	97.13
Image-5	99.96	99.27	98.51	97.24	97.06
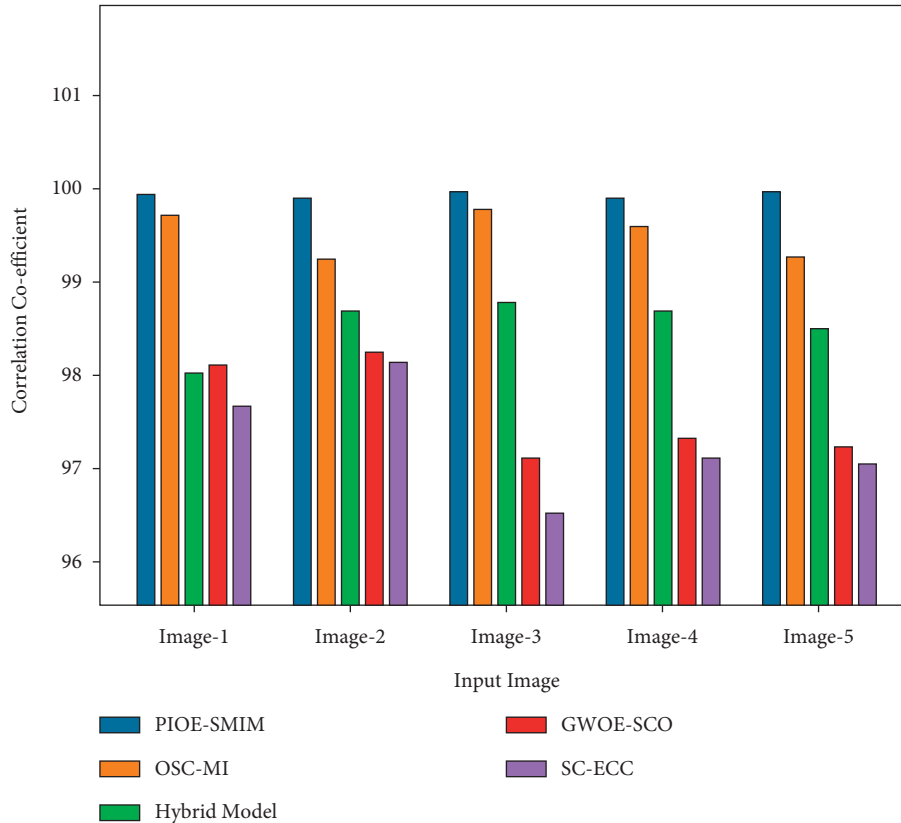
FIGURE 7: CC analysis of PIOE-SMIM technique with existing approaches.

TABLE 6: Result analysis of PIOE-SMIM technique with recent approaches with respect to PSNR (with and without attack).

Test images	PIOE-SMIM		OSC-MI		Hybrid model		GWOE-SCO		SC-ECC	
	Attack	WOA	Attack	WOA	Attack	WOA	Attack	WOA	Attack	WOA
Image-1	58.43	59.37	56.93	57.92	55.91	57.15	55.39	56.31	45.53	46.53
Image-2	58.99	59.99	55.97	57.04	55.61	56.81	55.46	56.43	44.20	45.24
Image-3	57.86	58.82	57.27	57.90	55.71	57.10	55.28	56.69	44.18	45.07
Image-4	59.37	60.28	58.16	59.04	55.21	56.31	54.27	55.23	44.40	45.70
Image-5	58.52	59.53	57.96	59.04	55.68	57.06	54.76	56.25	43.40	44.78

have gained decreased CC of 99.27, 98.51, 97.24, and 97.06, correspondingly [34].

Table 6 reports a detailed study of the PIOE-SMIM model with recent models under the existence and non-existence of attacks. Figure 8 showcases the PSNR examination of the PIOE-SMIM model with other models under the existence of attacks. The results indicated that the PIOE-SMIM model has gained effectual outcomes with increased PSNR. For instance, with image 1, the

PIOE-SMIM model has showcased an increased PSNR of 58.43 dB whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have depicted reduced PSNR of 56.93 dB, 55.91 dB, 55.39 dB, and 45.53 dB respectively [35]. Similarly, with image 5, the PIOE-SMIM model has reported an improved PSNR of 58.52 dB whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have exhibited lower PSNR of 57.96 dB, 55.68 dB, 54.76 dB, and 43.40 dB, respectively.
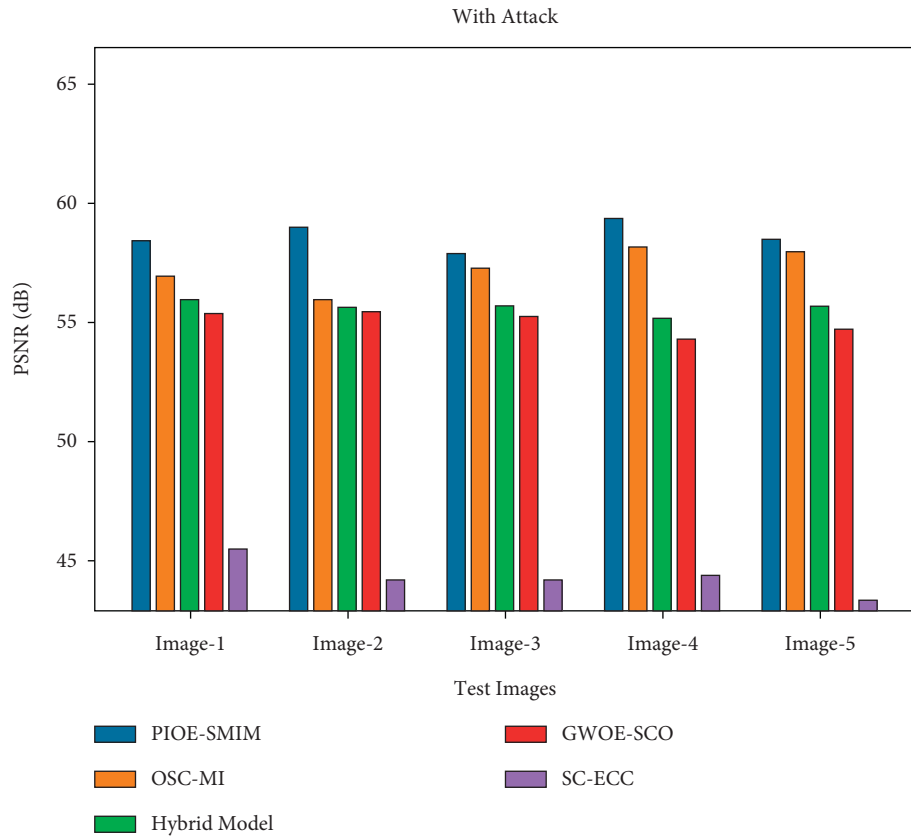
FIGURE 8: PSNR analysis of PIOE-SMIM technique with attacks.
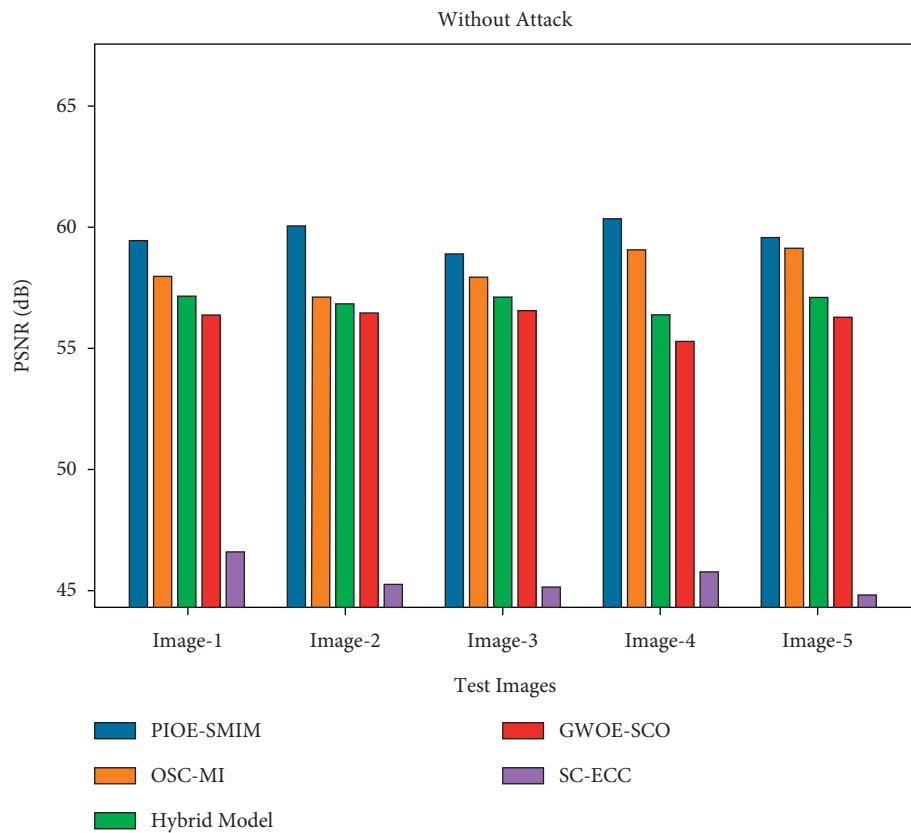
FIGURE 9: PSNR analysis of PIOE-SMIM technique without attacks.

Figure 9 illustrates the PSNR inspection of the PIOE-SMIM model with other models without attacks (WOA). The results designated that the PIOE-SMIM model has expanded capable outcomes with amplified PSNR. For instance, with image 1, the PIOE-SMIM model has showcased an improved PSNR of 59.37 dB whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have portrayed compact PSNR of 57.92 dB, 57.15 dB, 56.31 dB, and 46.53 dB, respectively. Likewise, with image 5, the PIOE-SMIM model has stated enhanced PSNR of 59.53 dB whereas the OSC-MI, hybrid, GWOE-SCO, and SC-ECC models have displayed inferior PSNR of 59.04 dB, 57.06 dB, 56.25 dB, and 44.78 dB, respectively.

The abovementioned tables and figures implied that the proposed PIOE-SSIM model has accomplished enhanced security in medical image transmission. Therefore, the proposed PIOE-SSIM model can be employed in the real-time hospitals and healthcare institutions to securely transmit medical images.

5. Conclusion

In this article, a new PIOE-SMIM model has been developed for accomplishing a secure image transmission system. The proposed PIOE-SMIM technique mainly focuses on the design of SSC and the encryption process. At the initial stage, the medical images are converted into a collection of 12 shares using the SSC approach. In addition, the ECC scheme is employed for the encryption process. In order to optimize key generation process in the ECC model, the PIO algorithm is exploited to maximize PSNR. Finally, on the receiver side, the decryption and share reconstruction processes are performed to construct the original images. For demonstrating an enhanced performance of the PIOE-SMIM technique, a comprehensive simulation study is made and the results highlighted the supremacy of the PIOE-SMIM model over other techniques. In the future, blockchain technology can be applied to improve security in the healthcare sector. Blockchain is applied in healthcare and other industries. In healthcare, a Blockchain network stores and shares patient data among hospitals, labs, pharmacies, and clinicians. Blockchain can detect deadly medical blunders. This could improve medical data sharing's performance, security, and transparency. This tech analyses medical records. Blockchain can improve healthcare and other industries. Blockchain can uncover hazardous medical errors. Blockchain avoids dishonesty to enhance clinical trial results.

Data Availability

The manuscript contains all of the data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Zhang, B. Rahmatullah, S. Wang, A. Zaidan, and P. Liu, "A review of research on medical image confidentiality related technology coherent taxonomy, motivations, open challenges and recommendations," *Multimedia Tools and Applications*, 2020.

[2] W. El-Shafai, F. Khallaf, E. El-Rabaie, and F. El-Samie, "Robust medical image encryption based on DNA-chaos cryptosystem for secure telemedicine and healthcare applications," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 10, pp. 9007–9035, 2021.

[3] Y. Tan, J. Qin, L. Tan, H. Tang, and X. Xiang, "A survey on the new development of medical image security algorithms, Cloud Computing and Security," in *Proceedings of the International Conference on Cloud Computing and Security*, pp. 458–467, Okinawa Japan, September 2018.

[4] K. Balasamy and S. Suganyadevi, "A fuzzy based ROI selection for encryption and watermarking in medical image using DWT and SVD," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7167–7186, 2021.

[5] Z. Mishra and B. Acharya, "High throughput and low area architectures of secure IoT algorithm for medical image encryption," *Journal of Information Security and Applications*, vol. 53, Article ID 102533, 2020.

[6] N. Zermi, A. Khaldi, R. Kafi, F. Kahlessenane, and S. Euschi, "A DWT-SVD based robust digital watermarking for medical image security," *Forensic Science International*, vol. 320, Article ID 110691, 2021.

[7] D. Ravichandran, A. Banu S, B. K. Murthy, V. Balasubramanian, S. Fathima, and R. Amirtharajan, "An efficient medical image encryption using hybrid DNA computing and chaos in transform domain," *Medical, & Biological Engineering & Computing*, vol. 59, no. 3, pp. 589–605, 2021.

[8] S. Kumar, B. Panna, and R. K. Jha, "Medical image encryption using fractional discrete cosine transform with chaotic function," *Medical, & Biological Engineering & Computing*, vol. 57, no. 11, pp. 2517–2533, 2019.

[9] M. Gafsi, N. Abbassi, M. A. Hajjaji, J. Malek, and A. Mtibaa, "Improved Chaos-Based Cryptosystem for Medical Image Encryption and Decryption," *Scientific Programming*, vol. 2020, pp. 1–22, Article ID 6612390, 2020.

[10] U. Gowshika and T. Ravichandran, "A smart device integrated with an android for alerting a person's health condition: internet of things," *Indian Journal of Science and Technology*, vol. 9, no. 6, 2016.

[11] B. Jaishankar, V. Santosh, S. Aditya Kumar, P. Ibrahim, and N. Arulkumar, "Blockchain for securing healthcare data using squirrel search optimization algorithm," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1815–1829, 2022.

[12] M. Susmita and P. Mohan, "Digital mammogram inferencing system using intuitionistic fuzzy theory," *Computer Systems Science and Engineering*, vol. 41, no. 3, pp. 1099–1115, 2022.

[13] S. Sambit, D. Sanchali, and D. Swapan, "A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA," *The Journal of Supercomputing*, vol. 76, no. 8, pp. 5849–5861, 2020.

[14] M. K. Hasan, S. Islam, R. Sulaiman et al., "Lightweight encryption technique to enhance medical image security on internet of medical things applications," *IEEE Access*, vol. 9, pp. 47731–47742, 2021.

[15] P. T. Akkasaligar and S. Biradar, "Selective medical image encryption using DNA cryptography," *Information Security Journal: A Global Perspective*, vol. 29, no. 2, pp. 91–101, 2020.

[16] W. Song, C. Fu, Y. Zheng, L. Cao, and M. Tie, "A practical medical image cryptosystem with parallel acceleration," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–15, 2022.

[17] B. D. Parameshachari and H. T. Panduranga, "Medical image encryption using SCAN technique and chaotic tent map system," in *Recent Advances in Artificial Intelligence and Data Engineering*, pp. 181–193, Springer, Singapore, 2022.

[18] P. Rashmi, M. C. Supriya, and Q. Hua, "Enhanced Lorenz-Chaotic Encryption Method for Partial Medical Image Encryption and Data Hiding in Big Data Healthcare," *Security and Communication Networks*, vol. 2022, Article ID 9363377, 2022.

[19] X. Wang, S. Yin, M. Shafiq et al., "A New V-Net Convolutional Neural Network Based on Four-Dimensional Hyperchaotic System for Medical Image Encryption," *Security and Communication Networks*, vol. 2022, Article ID 4260804, 2022.

[20] F. Masood, M. Driss, W. Boulila et al., "A lightweight chaos-based medical image encryption scheme using random shuffling and XOR operations," *Wireless Personal Communications*, pp. 1–28, 2021.

[21] M. Duraisamy and S. Balamurugan, "Multiple share creation scheme with optimal key generation for secure medical image transmission in internet of things environment," *International Journal of Electronic Healthcare*, vol. 11, no. 1, p. 1, 2021.

[22] A. Hafsa, A. Sghaier, J. Malek, and M. Machhout, "Image encryption method based on improved ECC and modified AES algorithm," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 19769–19801, 2021.

[23] Z. Cui, J. Zhang, Y. Wang et al., "A pigeon-inspired optimization algorithm for many-objective optimization problems," *Science China Information Sciences*, vol. 62, no. 7, 2019.

[24] R. R. Bhukya, B. M. Hardas, T. Ch et al., "An automated word embedding with parameter tuned model for web crawling," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1617–1632, 2022.

[25] D. K. Jain, S. K. S. Tyagi, M. Prakash, and L. Natrayan, "Metaheuristic optimization-based resource allocation technique for cybertwin-driven 6G on IoE environment," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4884–4892, 2022.

[26] S. Rajendran and U. M. Sankareswaran, "A novel pigeon inspired optimization in ovarian cyst detection," *Current Medical Imaging Reviews*, vol. 12, no. 1, pp. 43–49, 2016.

[27] P. Karamjeet Singh and P. Singh, "Image encryption and decryption using blowfish algorithm in Matlab," *International Journal of Scientific Engineering and Research*, vol. 4, no. 7, 2013.

[28] S. Ka car, I. Pehlivan, and A. Zengin, "Secure image encryption algorithm design with a novel chaos-based s-box," *Chaos, Solitons & Fractals*, vol. 95, pp. 92–101, 2017.

[29] K. S. Kannan, K. G. Suma, and S. Neelakandan, "A secured healthcare medical system using blockchain technology," in *Lecture Notes in Electrical Engineering*vol. 828, Singapore, Springer, 2022.

[30] G. L. N Murthy, E. Fantin Irudaya Raj, and N. Arulkumar, "Blockchain with deep learning-enabled secure healthcare data transmission and diagnostic model," *International Journal of Modeling, Simulation, and Scientific Computing*, .

[31] M. Ranjith Kumar, V. Chandra Shekhar Rao, R. Anand, and H. Singh, "Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm," *Measurement*, vol. 183, 2021.

[32] P. P. Mathai, C. Karthikeyan, T. Kavitha et al., "Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, 2021.

[33] C. Al-Atroshi, V. K. Nassa, B. Geetha et al., "Deep learning-based skin lesion diagnosis model using dermoscopic images," *Intelligent Automation & Soft Computing*, vol. 31, no. 1, pp. 621–634, 2022.

[34] G. Sunitha and K. Geetha, "Aditya Kumar Singh Pundir, Intelligent deep learning based ethnicity recognition and classification using facial images," *Image and Vision Computing*, vol. 121, 2022.

[35] A. Harshavardhan, Prasanthi boyapati, alhassan alolo abdul-rasheed akeji, aditya kumar singh pundir, and ranjan walia, "LSGDM with biogeography-based optimization (BBO) model for healthcare applications," *Journal of Healthcare Engineering*, vol. 2022, 2022.

Retraction

Retracted: Ageing Design of Urban Park Landscape Based on Computer Virtual Simulation Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Fan, "Ageing Design of Urban Park Landscape Based on Computer Virtual Simulation Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3150371, 9 pages, 2022.

Research Article

Ageing Design of Urban Park Landscape Based on Computer Virtual Simulation Technology

Jing Fan [ID]

School of Art Design, Xi'an FanYi University, Xi'an 710105, Shaanxi, China

Correspondence should be addressed to Jing Fan; fanjing202204@163.com

Greenery are the parks in the urban areas and it becomes progressively most significant as cities grow more crowded. In the end, urban parks contribute to the population's healthiness and welfare by providing opportunities intended for physical and social activity, leisure, and relaxation. In order to construct "Digital Land," computer and system software and hardware were used. Computer simulations are used to examine the value of a city garden's landscape design. Digital models and multimedia performances are being built using computer-aided design (CAD), this underlines the need for digitizing data for landscape design. A region's landscapes can only be accurately assessed if sufficient information is available on the elements that influence the people's awareness of landscape quality, as well as the kind, method, and effective rate of each of them. We use an LVQANN (artificial neural network) to forecast the landscape aesthetic assessment of urban parks and priorities the model's significant factors in this research. User viewpoint and artificial neural network modelling were utilized in conjunction to assess the aesthetic quality of the urban park's environment. This was done for two reasons. The design of urban parks decision support system is known as MATLAB software's multilayer perceptions model, which gives the ability to anticipate landscape visual significance in innovative parks. In this study, the ANN LVQ model is used to execute an ageing design of an urban park landscape based on a computer virtual simulation application. An example land is selected as input and area linked to sunny spot, top view, and so on is fixed. The ANN tool box is used to develop this application in MATLAB 2018b software. The following approach is used to create the ideal urban park landscape model. An accuracy of 89.23 percent, a sensitivity of 87.34 percent, and a recall of 78.93 percent were achieved, outperforming the approach and competing with the current model.

1. Introduction

For inhabitants, urban parks provide low-cost chances to connect with nature in their everyday lives, while also improving the aesthetics of the city. In order to accurately analyze a region's landscapes, it is necessary to know the kind, method, and effect rate of each of the influential aspects, as well as how they affect the user's sense of landscape quality. Environmental, biological, social, cultural, and physiological aspects all have a role in landscape's aesthetic value. For the majority of individuals, landscape variety, kind of landscape, and people's taste and ideals all play a role in aesthetic choice.

Designing Tehran's urban parks now adheres to the guidelines outlined in Vice President for Strategic Planning

and Supervision publication No. 203, with some minor modifications. This has resulted in a dramatic decrease in the park's landscape structure's aesthetic value, and there is no specific framework for increasing landscape quality. Which landscape aspects are most appealing to users and which are the most important variables in influencing their view of the environment? How can we forecast the aesthetic worth of urban parks in the eyes of observers by constructing the most suitable structure? 'The factors that affect the aesthetic quality of natural surroundings were found to be the viewpoint, land slope, and landform. Researchers feel that assessing the visual excellence of a land and its ecosystem requires paying close attention to the physical conditions of the landscape (which is illustrated in Figure 1).

Figure 1: Urban park landscape.

This has looked at the role that water plays in enhancing the aesthetics of the park and drawing in more visitors. Water features like ponds, pools, runnels, fountains, etc., should be included into landscape design, says one researcher, in order to provide park visitors with more diversity and enjoyment. For the evaluation of urban green space, Jahani and Mohammadi Fazel employed CNN modelling to determine characteristics and indications such as bared surfaces, architectural features, water, trees and shrubs, sports equipment, and mountain vistas as beneficial landscape beautifying aspects. Landscapes are classed as high-quality if they offer access to water, leisure, and sports equipment, according to studies. In addition to trees and plants, architecturally appealing buildings, and mountain vistas, urban green space has a better landscape quality. With the introduction of a multilayer perception model, the diversity of tree species and pure natural regions devoid of human interventions such as buildings, paths, and park equipment were given a boost in the landscape quality of natural areas. It is possible to utilize any of the design concepts outlined in these studies as indications of the current research as shown in Figure 1.

The contribution of the work is as follows:

(i) To execute an ageing design of an urban park landscape by using the ANN LVQ model based on a computer virtual simulation application.

(ii) To select the example land and taking it as an input and area linked to sunny spot, top view, and so on is fixed and perform evaluation.

1.1. Literature Survey. Jo and Jeon consider the relevance of human components; we determined that designs are required to encourage an adequate degree of park-goers to boost their overall contentment with the surroundings. An urban park setting may be made more enjoyable by using this study's results as design recommendations and supporting data [1].

Yu et al. As a result of using ArcGIS 3-D analysis, we want to present a novel study viewpoint and a service that provides techniques for creating layouts of such places that are beneficial for the elderly in order to better assure urban sustainability civilization [2].

Li and Xu Moreover, the research study and effectiveness test results show that the realized distribution rationality technology is better than the competition. This method might serve as a scientific reference and basis for the present landscape gardening industry to include 3-D images and other cutting-edge technologies [3].

Van Vliet et al. In an urban park, the natural components and a wide range of floral species are vital while the services provided vary widely in value among various sectors of the public Research like this might help park planners and legislators make better choices. As a result, it shows the value of establishing a virtual park in environmental preference study. Check out the full-length article [4].

Luo and Liu. As the rehabilitation environment evolves, this study examines present empirical research methodologies from the perspectives of subject and object, pointing out their flaws, forecasting their future development trends, and ultimately proposing future empirical research. The study should be more rigorous and methodical, with a better knowledge of what the research is about, and the establishment of an assessment system with the most practical significance possible should be encouraged [5].

Yang The urban parks landscape design and analysis of city parks significance may be improved by using computer simulation. More and more people are realising how important digitising information is to creating digital models and multimedia performances in computer-aided design (CAD), particularly when it comes to the landscape design process [6].

Ma et al. Analyzing the garden's rational distribution is based on a new three-dimensional picture. Virtual reality technology was used to build a three-dimensional simulation of the distribution landscape's efficacy. Finally, the garden, rational medication delivery methods, case studies, benefits, and results of performance tests will be discussed [7].

Irvine et al. To better understand the challenges of combining multidisciplinary visioning and modelling of green space design and performance evaluation in the Thai town of Cha Am, we interviewed students from three universities: Thammasat University, Mahidol University, and AIT. Examples from Phnom Penh illustrate how 3D visualisation and modelling might help people better comprehend the benefits of green space. For urban green space design to be successful, we believe that a decision support system that is easy to use is essential [8].

Bishop and Rohrmann findings from the study suggest that the most meticulous and time-consuming computer simulations may not always reflect actual-world behaviour. However, the realism evaluations of the viewers were largely positive, and the contrasts between day and night circumstances in the simulated environment are almost identical to those in the actual world. The results reveal areas in need of improvement and reassessment [9].

Jeon and Hong. Various human activities in parks have been shown to have a significant effect in influencing soundscape perception. There was also a strong link between how people perceived soundscapes and the aesthetic quality, simplicity, and feeling of enclosure provided by a given environment. As a consequence of this study, soundscapes in urban parks might benefit from the findings, even if they are only applicable to Seoul's specific settings [10].

Daniel and Meitner. Each degree of realism–abstraction was used to portray the identical woodland landscape images in this investigation. One similar set of forest landscape images was given to various groups of observers, and each group was asked to score the perceived scenic quality of the scenes. Computer-generated landscape visualisations may not be accurate representations of the landscape since correlations between judgments of the same landscapes in various visualisation situations were very low [11].

Lange. This investigation focuses on the subject of whether, how, and to what extent virtual landscapes can be used to accurately reproduce genuine landscapes seen in images. The settlements of Schwyz and Ingenbohl–Brunnen, located on the shores of Lake Lucerne in Central Switzerland, are the focus of this research [12].

Bishop and Lange. The field of visualisation technology has a lot to offer, and notable developers discuss the latest advancements in hardware, software, and display technology. Concerns about the legitimacy and ethics of the use of visual simulations are raised. Authors from forestry, agriculture, mining, energy, and urban planning have contributed cutting-edge applications. Finally, a look at what's to come in terms of augmented reality and public engagement is included in the volume's last chapter [13].

Bishop. The view analysis of Two-Dimensional GIS's approach is ineffective in circumstances when significant 3D aspects are present. For landscape visual interpretation, a new standard is required due to an increase in agent-based modelling. Both GIS-based visual modelling and 3D-based visual modelling are examined in this paper [14].

Thompson It is suggested that open space be defined and used in a more flexible manner, taking into account 'loosefit' landscapes that provide possibilities for the socially marginalised and environmentally moving within a dynamic context of urban structures and networks [15].

Wergles and Muhar. Visualizations, on the other hand, were more effective in conveying particular features of the design because of their capacity to draw the viewer's attention to focal points. As a means of conveying design concepts, visual representations may work well, but they must be tailored to the audience in order to be effective [16].

Sheppard Ethical and scientific validity may be improved by providing direction and assistance to visualisation practitioners, according to a suggested framework. An preliminary code of ethics is offered for evaluation, testing, and change by other researchers and users until more thorough results are gathered from the vast amount of study required on this topic. Such a code should incorporate general principles and guidelines for ethical behaviour in the creation, presentation, and analysis of visualisations, as well as the analysis of user comments [17].
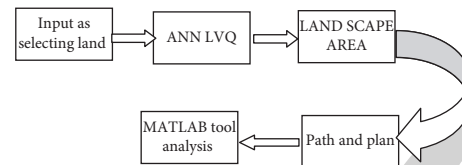
Figure 2: Proposed block.

Appleton and Lovett. According to the data, it seems that certain characteristics are more significant than others when it comes to realism. Ratings were significantly influenced by foreground vegetation and the general appearance of the ground surface, despite a large variation in ratings across the three situations evaluated [18].

Kort et al. On four measures, researchers discovered substantial variations. Environment and plants, on the other hand, did not seem to interact in a major way. Virtual environments have many characteristics with actual ones, but there are also significant distinctions. As a result, we believe that this is an important and fascinating topic for environmental psychology [19].

Orland et al. Using virtual reality (VR) technology to aid with environmental decision-making is the focus of this article. Human-computer interfaces are discussed and categorized, and then we examine how VR's qualities match up with landscape representation demands [20].

The drawback of the existing approaches is as follows, expensive direct observations, it requires numerous employs for monitoring the extended periods and restricted information only provided.

2. Methodology

In this section, urban park landscape-based analysis was performed using LVQ ANN technology. For multiclass classification, LVQ (learning vector quantization) is a prominent class of adaptable nearest prototypes classifiers. However, learning algorithms from this family have been developed on heuristic grounds so far. LVQ (Learning Vector Quantization) classifiers are reviewed in this article. A suggested taxonomy incorporates the most significant LVQ techniques to date. The fundamental elements of contemporary LVQ techniques are outlined in this article. 11- LVQ extractors are compared by means of a real-world dataset and 2 fake areas shown in Figure 2.

$$y = \{y(t)|t = o, 1, 2, \ldots d - 1\}. \tag{1}$$

Here, the quantity of Urban park area [5] for time. t, d: length of network traffic duration.

The domain represents a discrete-time series of finite length *according* to ANN theory. So,

$$x(n) = y(t), n = t, n = 0, 1, \ldots, d - 1. \tag{2}$$

This discrete series $x(n)$ represents the set given below,

$$\{x(0) = y(1), x(1) = y(2), \ldots .x(d - 1) = y(z)\}. \tag{3}$$
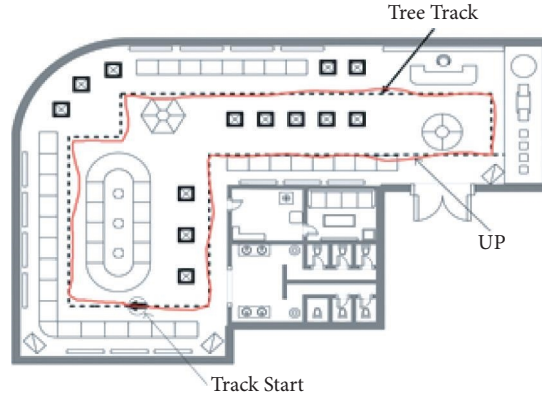
Figure 3: MATLAB-based ANN LVQ analysis.

Table 1: Parameter estimation.

Regression weights	Estimate	S.E.	C.R.	P
UP-1 <---URBAN PARK _SCAPE	1.000			
UP-2 <---URBAN PARK _SCAPE	1.440	0.149	9.675	***
UP-3 <---URBAN PARK _SCAPE	1.733	0.161	10.778	***
UP-4 <---URBAN PARK _SCAPE	1.894	0.174	10.875	***
UP-5 <---URBAN PARK _SCAPE	1.458	0.144	10.141	***
UP-6 <---URBAN PARK _SCAPE	1.246	0.126	9.886	***
UP-7 <---URBAN PARK _SCAPE	1.723	0.147	11.700	***
UP01-6 <---URBAN PARK _SCAPE_001	1.000			
UP01-5 <---URBAN PARK _SCAPE_001	3.596	1.396	2.576	0.010
UP01-4 <---URBAN PARK _SCAPE_001	4.129	1.584	2.607	0.009
UP01-3 <---URBAN PARK _SCAPE_001	4.297	1.648	2.607	0.009
UP01-2 <---URBAN PARK _SCAPE_001	0.721	0.491	1.469	0.142
UP01-1 <---URBAN PARK _SCAPE_001	1.403	0.631	2.223	0.026
URBAN PARK _SCAPE_002-6 <--- URBAN PARK	1.000			
URBAN PARK _SCAPE_002-5 <---URBAN PARK	0.911	0.116	7.883	***
URBAN PARK _SCAPE_002-4 <---URBAN PARK	0.939	0.110	8.533	***
URBAN PARK _SCAPE_002-3 <---URBAN PARK	1.151	0.122	9.441	***
URBAN PARK _SCAPE_002-2 <---URBAN PARK	0.827	0.092	9.006	***
URBAN PARK _SCAPE_002-1 <---URBAN PARK	1.042	0.105	9.954	***
TAD-6 <---LANDSCAPE analysis	1.000			
TAD-5 <---LANDSCAPE analysis	1.208	0.139	8.662	***
TAD-4 <---LANDSCAPE analysis	1.277	0.131	9.752	***
TAD-3 <---LANDSCAPE analysis	1.586	0.157	10.095	***
TAD-2 <---LANDSCAPE analysis	1.139	0.134	8.502	***
TAD-1 <---LANDSCAPE analysis	1.579	0.153	10.328	***
Essential service dept for eco-tourism < ---URBAN PARK _SCAPE	0.233	0.057	4.084	***
Essential service dept for eco-tourism < ---URBAN PARK _SCAPE_001	0.670	0.298	2.251	0.024
Essential service dept for eco-tourism < ---URBAN PARK	−0.051	0.152	−0.333	0.739
Essential service dept for eco-tourism < ---LANDSCAPE analysis	1.050	0.185	5.688	***

The ANN $x(n)$ with finite length d is considered as a sequence of periodic nature. Subsequently, with the theory of ANN.

$$X(k) = \sum_{n=0}^{d-1} \left(x(n) e^{-j(2\pi/d)nk} \right), \quad 0 \le k \le d-1. \tag{4}$$

Consequently, $x(n)$ is expanded as the equation given below:

$$x(n) = \frac{1}{z} \sum_{k=0}^{z-1} \left(Y(k) e^{j(2\pi/d)nk} \right), \quad 0 \le n \le d-1. \tag{5}$$

From (2) and (5), the value of $y(t)$ at time t is obtained as follows:

$$\bar{y}(t) = y(n) = \frac{1}{z} \sum_{k=0}^{z-1} \left(Y(k) e^{j(2\pi/d)nk} \right), t = n, 0 \le n \le d-1. \tag{6}$$

(a)

(b)

(c)

(d)

FIGURE 4: Top view for selected urban park landscape.

Since the network traffic considers $x(t) \geq 0$, the equation below is attained for accessing traffic in network.

$$\begin{cases} \tilde{y}(t) = \dfrac{1}{z} \sum_{k=0}^{z-1} \left(Y(k) e^{j(2\pi/d)nk} \right), & 0 \leq n \leq d-1, \\ \\ \tilde{y}(t) \geq 0. \end{cases} \tag{7}$$

Equation (7) represents that the traffic across network is extended to Fourier series with the duration length d as well as network traffic correlation at $d-1$ time points. The Fourier series expansion matrix is given by the following:

$$\hat{y} = \begin{bmatrix} \hat{y}_{0,0} & \hat{y}_{0,1} \cdots & \hat{y}_{0,d-1} \\ \hat{y}_{1,0} & \hat{y}_{1,1} \cdots & \hat{y}_{1,d-1} \\ \hat{y}_{d-1,0} & \hat{y}_{d-1,1} \cdots & \hat{y}_{d-1,d-1} \end{bmatrix}. \tag{8}$$

Hence, we get the following:

$$\hat{y} = \left(\hat{y}_0, \hat{y}_1, \ldots \hat{y}_{d-1} \right)^T. \tag{9}$$

Here, T: matrix transposition operation

$$\text{Consequently, } \hat{y}_i = \left(\hat{y}_{i,0}, \ldots \hat{y}_{i,d-1} \right). \tag{10}$$

In which, $i = 0, 1, \ldots \ldots, d-1$
From (7),

$$\hat{y}_{i,k} = X(k) e^{j(2\pi/d)ik}, \tag{11}$$

where. $i, k = 0, 1, \ldots, d-1$

Here, (8) captures the relation within the network traffic effectively. Subsequently,

$$\hat{y} = UDV^T, U, V, D \text{ and } V \text{ are } d \times d \text{ matrices.} \tag{12}$$

D and V represent the network traffic features and equation (12) captures the area dimensions features.

$$K(d) = \frac{p_0}{2} + \sum_{k=1}^{d} \left(p_k \cos kx + q_k \sin kx \right). \tag{13}$$

The authentication key for the time duration $d+1$ is given by,

(a)

(b)

(c)

(d)
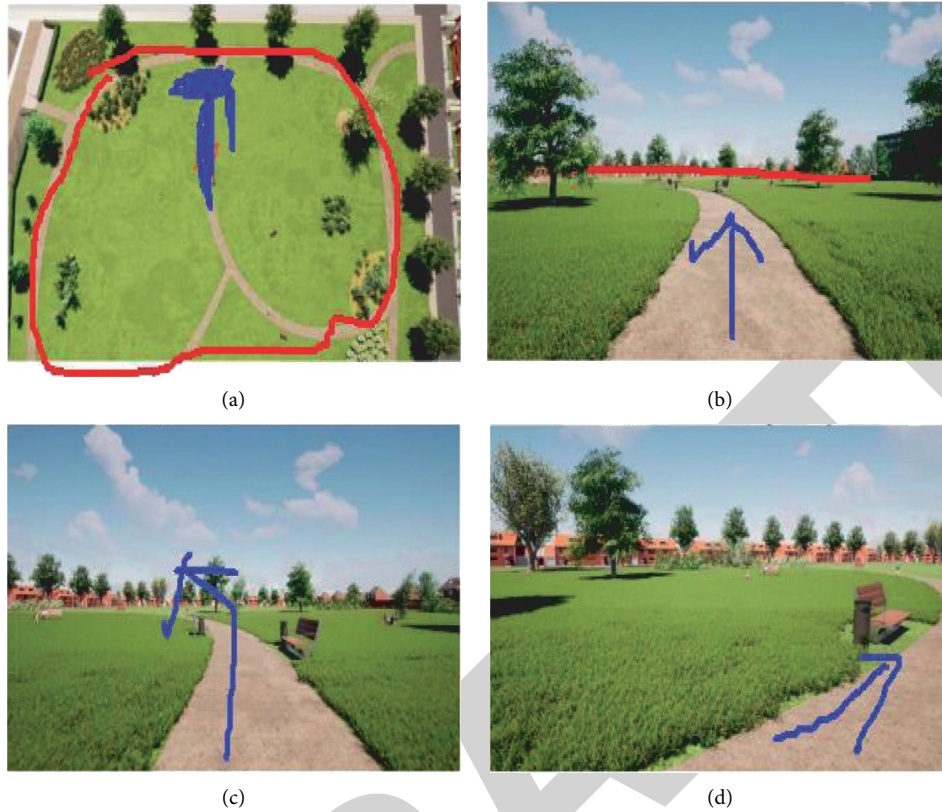
Figure 5: Computer level design with ANN.
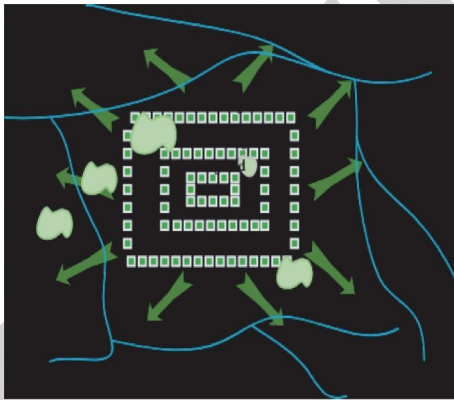
Figure 6: Inner layout.

$$K(d+1) = \frac{p_0}{2} + \sum_{k=1}^{d+1} \left(p_k \cos kx + q_k \sin kx \right), \tag{14}$$

$$= K(d) + \left(p_{d+1} \cos (d+1)x + q_{d+1} \sin (d+1)x \right). \tag{15}$$

The authentication key $K(d+1)$ for time period $d+1$ is calculated by $K(d)$ as well as the Fourier series coefficients p_{d+1} and q_{d+1}.

In statistical classification and prediction, LVQ (Learning Vector Quantization) is a family of methods that try to learn prototypes (codebook vectors) that represent class areas. Voronoi partitions are used to

define the class areas as hyperplanes between prototypes. Late in the '80s, Teuvo Kohonen came up with LVQ1 an algorithm that has since been used in many different ways. There has been a small but dedicated group of researchers studying LVQ algorithms since their introduction. "Learning Vector Quantization" or "LVQ" appeared in the titles or abstracts of 665 academic articles that were indexed by the ISI Web of Science in November 2013. This article presents a survey of recent developments in the subject.

In order to achieve competitive learning's, LVQ algorithms are similar to SOMs and c-means. During training, the winner-takes-all learning rule and variations based on iterative URBAN PARK _SCAPE_002ates to just selected components or neighborhoods are used to create competitive learning algorithms. Class-labeled prototypes are obtained via supervised learning in the original LVQ technique and most recent extensions (classifiers). Unsupervised learning may also be used to train LVQ for clustering without labels. We will exclusively cover LVQ classifiers in this work.

Data are supposed to have a Gaussian mixture probability density function (URBAN PARK _SCAPE_002f). Logarithmically maximizing the difference between the proper class's and erroneous class 'URBAN PARK _SCAPE_002f ratios is a cost function given a data point.

In addition, GLVQ and the original LVQ algorithms have been improved in terms of their initialization sensitivity [15].
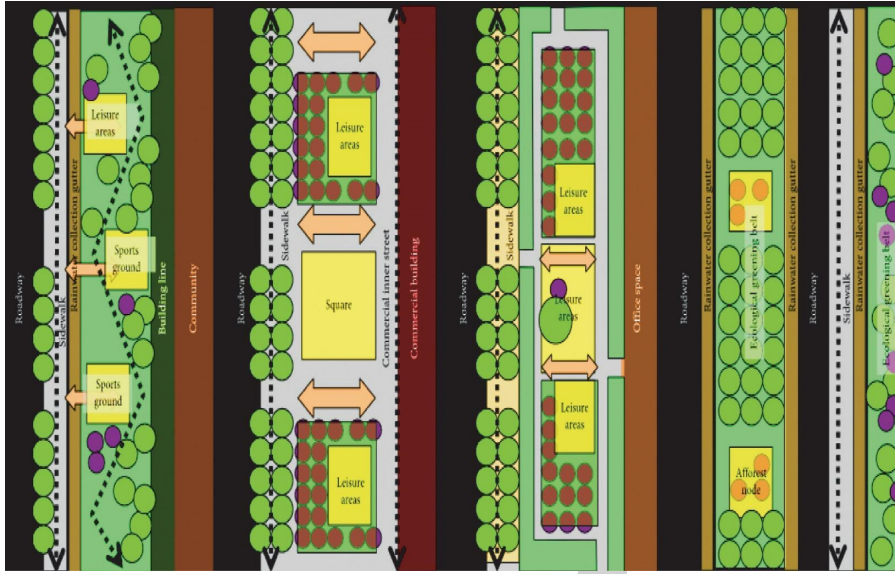
Figure 7: Final layout.

Table 2: Results comparison.

Models		LVQ [5]	ANN [6]	Deep ML [7]	Deep GA [7]	EST [30]	ANN_LVQ proposed
MATLAB	Accuracy	77.877	83.8687	87.6	84.3035	87.77	88.386
	Specificity	77.7373	83.743	83.7	83.8737	87.37	85.37
	Sensitivity	57.6473	74.43	87.77	87.3456	87.57	88.37
Theory	Accuracy	87.3374	83.76	83.45	83.6433	87.737	88.43
	Specificity	83.7378	83.56	83.7	83.874	85.74	86.37
	Sensitivity	87.6557	87.7	87.3	83	86.84	87.33

For example, LVQ may be used in image and signal processing, the biomedical area, and medicine and industry, just to mention a few examples of its many uses There is a vast collection of bibliographic data accessible in the system.

In this study, we give a thorough overview of the most significant supervised LVQ algorithms since Teuvo Kohonen's initial work. Our taxonomy and techniques for LVQ classification are presented here. A total of eleven different LVQ algorithms are tested on a variety of simulated and real-world datasets. Depending on the kind of datasets, we explore the benefits and drawbacks of each strategy.

Assuming that the input space contains a number of reference vectors Wk. Typically; a class is allocated a number of reference vectors. This means that input vector x has been assigned a classification based on the closest reference vector. We may think of Wk (t) as the discrete-time representation of Wk sequences. In this part, we will concentrate on LVQ2.1, which has been offered before. This is how LVQ2.1 URBAN PARK_SCAPE_002ates your reference vectors from a set of correctly specified beginning values.

$$\min\left(\frac{d_i}{d_j}, \frac{d_j}{d_i}\right) > S. \tag{16}$$

As in, di = Ix−WD and d_i = Ix−WD + wjl.

Based on the notion of pushing decision boundaries toward Bayesian bounds by applying attractive and repulsive pressures from x, LVQ2.1 implements an algorithmic approach (shown in Figure 3). As a result, the reference vectors diverge in the long term since no consideration is paid to the Wk location. Referencing the class distributions using multiple reference vectors has been suggested as a way of preventing reference vector divergence. However, LVQ3 has the same properties as the earlier LVQ2.1 and so does not address the issue of reference vector divergence.

3. Results and Discussion

In this section, a brief discussion of urban park landscapes are finding via the MATLAB-based ANN computer vision model.

The parameter estimation is shown in Table 1:

Figure 4 clearly explains about the top view model of a proposed design estimating image; in this, using ANN LVQ model attains proper landscape architecture.

Figure 5 clearly explains about different views of selected urban area, in this proposed green roof initiatives have been used to enhance climatic conditions, conserve and develop plant, encourage outdoor activities and active lifestyles, stimulate social connection and exchange, and create healthy urban circumstances for cognitive and emotional well-being.

Figure 6 clearly explains about the inner layout of urban park landscape analysis, in this all-road maps and paths are added accordingly.
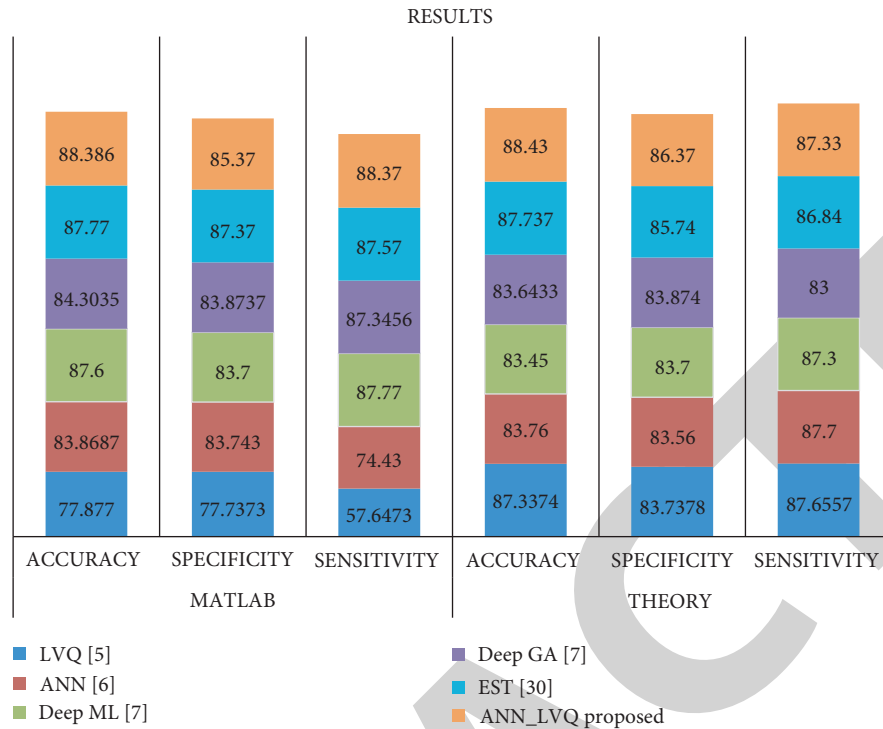
RESULTS

FIGURE 8: Comparison of results.

Given the limited extent of these green spaces, design standards must be simple: trees, bushes, and grass areas must be placed in such a manner that shaded and sunny places alternate; paved areas must also be given, with areas for playing and relaxing shown in Figure 7.

The results' comparison is shown in Table 2:

Figure 8 clearly explains about different comparison of results of proposed model, in this proposed ANN_LVQ model attains more improvement. Functionality variation arises in the process of simulation, this system compares the sensitivity of the system with random variations and this results in the decrease of computational speed.

4. Conclusion

In this research work, ageing design of urban park landscape based on computer virtual simulation application is implemented using ANN LVQ model. The visual quality of 100 urban park settings was assessed using citizen perception. A total of 15 landscape characteristics impacted the visual quality of the scene. According to the findings, the multi-layer perceptions model 15-8-1 with a configuration of 15-8-1 (15 input variables, 8 hidden layer neurons, and 1 output variable) has the best structure optimization performance. Three data sets were obtained with R2 values of 0.97, 0.88, and 0.90: training (validation), test, and (test). The visual quality of urban park landscapes is most affected by land slope and vegetation/building/hard surface ratios with model sensitivity values of 0, 24, 0, 7, and 0. In this an example land is selecting as input and fixing area related to sunny place, top view etc. This application is implemented on MATLAB 2018b software using ANN tool box. The

perfect urban park landscape model is designed using following technique. The measures like accuracy 89.23%, sensitivity 87.34% and recall 78.93% had been attained those are out performance the methodology and compete with present model.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] H. I. Jo and J. Y. Jeon, "Overall environmental assessment in urban parks: Modelling audio-visual interaction with a structural equation model based on soundscape and landscape indices," *Building and Environment*, vol. 204, 2021.

[2] W. Yu, H. Hu, and B. Sun, "Elderly Suitability of Park Recreational Space Layout Based on Visual Landscape Evaluation," *Sustainability*, vol. 13, no. 11, p. 6443, 2021.

[3] R. Li and D. Xu, "Distribution of landscape architecture based on 3D images and virtual reality rationality study," *IEEE Access*, vol. 8, pp. 140161–140170, 2020.

[4] E. van Vliet, G. Dane, M. Weijs-Perrée et al., "The influence of urban park attributes on user preferences: Evaluation of virtual parks in an online stated-choice experiment," *International journal of environmental research and public health*, vol. 18, no. 1, p. 212, 2020.

[5] J. Luo and H. Liu, "Analysis of Empirical Research Methods of Rehabilitation Landscape," *IOP Conference Series: Earth and*

Research Article

Vibration and Trajectory Tracking Control of Engineering Mechanical Arm Based on Neural Network

Xinjun Lei [1,2] **and Yunxin Wu** [1,3,4]

[1]*State Key Laboratory of High Performance Complex Manufacturing, Central South University, Changsha 410083, China*
[2]*SANY Automobile Manufacturing Co.,Ltd, Changsha 410100, China*
[3]*College of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China*
[4]*Light Alloy Research Institute, Central South University, Changsha 410083, China*

Correspondence should be addressed to Xinjun Lei; xinjun.l@csu.edu.cn

We offer a neural network-based control method to control the vibration of the engineering mechanical arm and the trajectory in order to solve the problem of large errors in tracking the path when the engineering mechanical arm is unstable and under the influence of the outside world. A mechanical arm network is used to perform tasks related to learning the unknown dynamic properties of a engineering mechanical arms keyboard without the need for prior learning. Given the dynamic equations of the engineering mechanical arm, the dynamic properties of the mechanical arm were studied using a positive feedback network. The adaptive neural network management system was developed, and the stability and integrity of the closed-loop system were proved by Lyapunov's function. Engineering mechanical arm motion trajectory control errors were modeled and validated in the Matlab/Simulink environment. The simulation results show that the management of the adaptive neural network is able to better control the desired path of the engineering mechanical arm in the presence of external interference, and the fluctuation range of input torque is small. The PID control has a large error in the expected trajectory tracking of the engineering mechanical arm, the fluctuation range of the input torque is as high as 20, and the jitter phenomenon is more serious. The use of detailed comparisons and adaptive neural network monitoring can perform well in manipulating the trajectory of the engineering mechanical arm. The engineering mechanical arm uses an adaptive neural network control method, in which the control precision of engineering mechanical arm motion trajectory can be improved and the out-of-control phenomenon of mechanical arm motion can be reduced.

1. Introduction

With the development of science and technology, machine technology is more and more applied in various fields of national life, which puts forward higher requirements for the dynamic performance of machine control system. Machine modeling and control is the main content of machine research and plays a decisive role in the development and application of machines. The machine control object has one or more multi-joint mechanical arms. Its static and dynamic models are extremely complex, with multiple inputs and outputs, which are combined with each other. The system is also closely related to its kinematics and dynamics and is

highly nonlinear [1]. During the operation of the machine control system, the system load changes greatly with time and has strong time variability. This kind of complex control system has high requirements on the model of the control object. How to improve the accuracy of modeling is the focus of current research. Many emerging modeling strategies are introduced, which provide a good foundation for improving control performance. Modern industry and engineering construction have higher and higher requirements for mechanical arms, which are reflected in the working speed and precision of the machine [2]. Its appearance makes it inevitable to apply more advanced modern control theory to the field of machine control to solve the problem of

nonlinear severe disaster control. Typical advanced modern control theories and technologies include rational control, adaptive control, neural network control, sliding mode variable structure control, and so on. Track control plays an important role in mechanical arm traffic control. In the actual use of the mechanical arm, the system must not only have perfect functions but also have good performance. The first means that the machine must work within a relatively wide range, and the last means that the machine must meet human needs. As a typical representative of engineering machines, mechanical arms have the kinematics and dynamics of general machines, that is, multi-variable, time-varying, random, and extremely nonlinear. The existence of uncertain factors significantly affects the performance of the mechanical arm control system, which is inevitable. Traditional object-based control methods are increasingly difficult to meet the control requirements of mechanical arm trajectory [3, 4].

The development of intelligent management theory makes neural network technology an effective way to solve optimal problems. The reason why the neural network can solve the strong nonlinear mapping problem is determined by its self-organization and self-learning ability. Neural networks can approximate any nonlinear function and have good adaptability to data mapping, which cannot be satisfied by traditional numerical calculation methods [5]. Neural networks are widely used in various areas of automatic control, especially system recognition, nonlinear system control, fault diagnosis, and fault tolerance monitoring, and play a very important role. Nerve networks play a strong role in the recognition of nonlinear systems. As a model for recognizing nonlinear systems, it provides a general model for recognizing neural networks. Therefore, it is the physical realization of the actual system and is suitable for online control. RBFNN based on RBF (radial basis function) is a typical transmission network. It is based on function approximation theory. Network learning is basically to find the optimal plane for data alignment to meet the requirements of multi-dimensional information training space. Any basis function contained in this best-fit plane consists of the latent layer neuron transfer function of the RBFNN. The complexity of mechanical arm kinematics and dynamics makes neural networks very useful and highly applicable in the field of mechanical arm control. The application of neural network technology in mechanical arm control can solve the problem of real-time online control of mechanical arms. On the one hand, by learning the sample data during the actual operation of the mechanical arm, the neural network can obtain the nonlinear relationship between the kinematics or dynamics input and output of the mechanical arm, which is usually represented by a nonlinear expression. The obtained nonlinear relationship can make the system immune to the influence of parameter uncertainty on the control because its establishment process does not directly depend on the actual parameters of the mechanical arm. On the other hand, the neural network control system is robust, which reduces its sensitivity to the system characteristics or parameter distortion of the mechanical arm, making the mechanical arm more intelligent, so as to achieve the purpose of real-time

control of the mechanical arm. Neural network is a relatively systematic and accurate control method in the trajectory tracking control of mechanical arm, which provides an effective way to study the modeling and control of mechanical arm. It can ensure that the position and velocity tracking errors of the end effector of the mechanical arm gradually converge to zero [6, 7]. Aiming at the problems of closed-chain structure, unknown dynamic parameters, and strong coupling in coordinated multi-mechanical arm systems, a position-force hybrid control strategy was designed based on fuzzy neural network. The whole system is modeled dynamically in the task space, and the idea of position-force hybrid control is adopted. The controllers are designed in the position space and the force space, respectively. By introducing the coordinated control amount into the position controller, the engineering mechanical arm is reduced in size.

2. Related Works

Neural network originated in the early 1940s and has experienced nearly 20 years of development. It was not until the late 1950s that the neural network theory was initially formed. For the problem of mechanical arm modeling and control, many researchers have done a lot of research and successfully verified it in experiments or practice. Schwartz-Leyzac and others proposed an online identification method of mechanical arm dynamic model based on RBFNN and achieved good results on 6-DOF industrial mechanical arm [8]. RBF and MLP (multi-layer perceptron) have been successfully used in 6-DOF series mechanical arms, offering a mechanical armic reverse kinematic solution based on a combination of neural networks [9]. Carney and others studied a new method of adaptive control of neural networks that has been studied and successfully used to control the movement of the PUMA560 mechanical arm [10]. Hsu and others offered a hybrid learning algorithm that combines neural network-based Kalman filters and BP algorithms, combined neural networks with traditional PID control used in mechanical armic adaptive control systems, and explored effective heuristic learning algorithms [11]. Abbas and others made efforts in the study of manipulative neural network management, and a model of adaptive neural network management of mechanical arms based on dynamic return networks was proposed [12]. A stable mechanical armic neural network controller was studied, and a control circuit consisting of a neural network management and control controller was proposed [13]. Patocka and others put forward an adaptive neural slip mode control method to track the path of uncertain mechanical arms, which effectively performs rapid control of the mechanical arm's path under system uncertainty [14]. Liu and others proposed a kinematic model of a mechanical arm that effectively solves the problem of mechanical arm control by offering a dynamic repetitive neural network with state delay input [15]. Blacker and others proposed a solution method of forward and inverse kinematics of mechanical arm based on RBFNN from the perspective of geometry and mathematics, which effectively avoided the tediousness of traditional mechanical

arm kinematics solution methods and provided convenience for the research of mechanical arm dynamics control [16]. Fu and others recommended RBFNN-based iteration training controller, which effectively solves the problem of low wristband speed when using iterative training control to control the mechanical arm track [17]. Liu and others put forward a hybrid controller consisting of an adaptive RBFNN and a PD controller, which effectively solves the powerful control problem of controlling the path of an unknown mechanical arm in the event of an external failure [18].

Based on this research, the invention provides a vibration and trajectory control engineering manipulator based on neural network. Starting from the kinematics and dynamics of engineering mechanical arm, this paper studies the RBFNN learning algorithm based on EC-RBF, applies it to the mechanical arm motion control algorithm, realizes the neural network solution of the inverse kinematics of mechanical arm, and designs the mechanical arm NNMARC based on the learning algorithm. The neural network learning method combining offline learning and online adjustment is used to identify the dynamic model of the mechanical arm. Finally, the control of the engineering mechanical arm track is realized.

3. Research Methods

3.1. Dynamic Characteristics of Engineering Mechanical Arm. The engineering mechanical arm is a multi-link mechanism with an open chain, the main body is fixed on the construction machinery, and an actuator is installed at the free end to suit the needs of the work. The two joints are connected by a rotating joint or a moving joint, and the interaction of the joints involves the movement of the joint and the movement of the mechanical arm to reach a different position. The kinematic equation of the engineering mechanical arm is the conversion of the mechanical arm end effect between the joint space and the working space. The dynamic equation represents the nonlinear relationship between the torque applied to each joint of the engineering mechanical arm and the angular displacement and angular velocity of each joint, that is, the movement of the engineering mechanical arm is controlled by controlling the torque. In order to achieve the goal of real-time monitoring, it is necessary to create a dynamic model of the mechanical arm and effectively monitor it. The main content of the study of the kinematics of the engineering mechanical arm is the creation of a coordinate system of the coordinates of the position of each link, the direction, and the relationship between the transitions between the different coordinates. The kinematics of a engineering mechanical arm has two problems, one is the forward kinematic problem and the other is the reverse kinematic problem, called reverse kinematics. The forward kinematics of the mechanical arm depends on the spatial position and orientation of the mechanical arm, taking into account each of the joint variables of the engineering mechanical arm. Reverse kinematics is the solution of each joint variable according to the spatial position and orientation of the engineering

mechanical arm. Both the front kinematic policy of the mechanical arm and the problem of reverse kinematics can be considered as problems of nonlinear mapping of the engineering mechanical arm between the joint distance and the work area [19].

For the engineering mechanical arm with n degrees of freedom, the connection position variable is represented by $q = [q_1, q_2, \ldots, q_n]^T$. It is assumed that the research object can be represented by m variables $x = [x_1, x_2, \ldots, x_m]^T$ $(m < n)$, and $x = f(q)$, where f is the forward kinematics equation. After the quadratic differentiation of X, we can get

$$x'' = J'(q)q' + J(q)q'', \tag{1}$$

where $J(q) = \partial f(q)/\partial q$ and $J(q) \in R^{m \times n}$ represents the Jacobian matrix of the mechanical arm.

The pseudo-inverse matrix (represented by $\overset{+}{J}(q)$) of the Jacobian matrix (JQ) of the mechanical arm is defined as follows:

$$\overset{+}{J} = \overset{T}{J}(J\overset{T}{J})^{-1}, \tag{2}$$

where $\overset{+}{JJ} = I_m$ and $I_m \in R^{m \times n}$ is the identity matrix.

For the engineering mechanical arm with n connecting pairs, the rotation direct drive can be expressed by the following relationship:

$$M(q)q'' + V_m(q, q') + G(q) + F(q') = \tau, \tag{3}$$

where $M(q) \in R^{n \times n}$ is the inertia matrix; $V_m(q, q')$ is the centripetal Coriolis force matrix; $G(q) \in R^n$ is the gravity matrix; $E(q')$ is the friction matrix; and $\tau \in R^n$ is the torque input vector.

Engineering mechanical arm dynamics refer to the study of the relationship between the driving force of each joint and the motion of the joint. The engineering mechanical arm is composed of multiple joints and multiple connecting rods.

Each degree of freedom is driven by a separate actuator. From the control point of view, the engineering mechanical arm control system itself is an automatic control system, which has the characteristics of multi-input and multi-output, nonlinearity, redundancy, and combination. The combination of dynamics is reflected in the relationship between the joint motion of the engineering mechanical arm and the torque applied to each joint. Therefore, the motion control task of the engineering mechanical arm is essentially a dynamic task. There is a strong combination between them, and the parameters are time-varying, and the model is also uncertain. The purpose of studying the dynamics of engineering mechanical arm is to control. In order to realize the accurate control of the motion of the mechanical arm, its dynamic model must be established. Dynamic model identification is one of the most basic but relatively difficult problems in the research of engineering mechanical arm control. There are two main methods to analyze the dynamic model of engineering mechanical arm.

3.1.1. Newton–Euler Equation. This method is based on the dynamic balance method of force, which can eliminate the internal force between joints while obtaining the motion

speed of the mechanical arm. This method can be used to analyze the general simple engineering mechanical arm system. For more complex systems, this method will appear extremely complex, so it is not applicable.

3.1.2. Lagrange Equation. It is the main equation of Lagrangian mechanics, which can be used to describe the motion of objects, especially for the study of theoretical physics. The Lagrange equation is functionally equivalent to Newton's second law in Newtonian mechanics. This method is based on Lagrange function balance method. In the establishment process, it is not necessary to obtain the internal force between the joints of the mechanical arm. The derivation is simple and clear, and the physical meaning of the parameters is clear. It is a method with strong applicability. This section will focus on the establishment of the dynamic model of the engineering mechanical arm based on this method.

3.2. Feedforward Neural Network. The neural network is the simplest type of neural network. Each neuron is located in a layer, and each neuron communicates only with neurons in the previous layer. The output of the previous layer is received and passed to the next layer without feedback between the layers. It is one of the most widely used and fastest growing neural networks. Research began in the 1960s and reached a very high level of theoretical research and practical application. The feedforward neural network (FNN) is a type of artificial neural network called the feeder network. The transmission neural network uses a unilateral multi-layer structure. Each layer contains several neurons. In this type of neural network, each neuron can receive signals from nerve cells in the previous layer and output them to the next layer. Layer 0 is called the input layer, the last layer is called the output layer, and the other intermediate layer is called the hidden layer. The hidden layer can be a single layer. It can also be a multi-layer conductive neural network with n input units, m output units, and N latent level units, as shown in Figure 1.

The output vector Z can be determined by the input vector X through the following formula:

$$z_i = \sum_{j=1}^{N} \left[\omega_{ij} \sigma \left(\sum_{k=1}^{n} v_{jk} y_k + \theta_{vj} \right) + \theta_{\omega i} \right], \quad i = 1, 2, \ldots, m,$$

$$(4)$$

where $\sigma(.)$ is the hidden layer neuron activation function, $\sigma(y) = 1/(1 + e^{-y})$; v_{jk} is the input layer to hidden layer interconnection weight; ω_{ij} is the weight of the connection between the hidden layer and the output layer; and θ_{vj}, $\theta_{\omega i}$ represent deviation weight.

By substituting the neural network weight values v_{jk} and ω_{ij} into the weight matrices V^T and W^T, respectively, the corresponding vectors of the neural network equation can be obtained as follows:

$$z = W^T \sigma (V^T y),$$

$$(5)$$
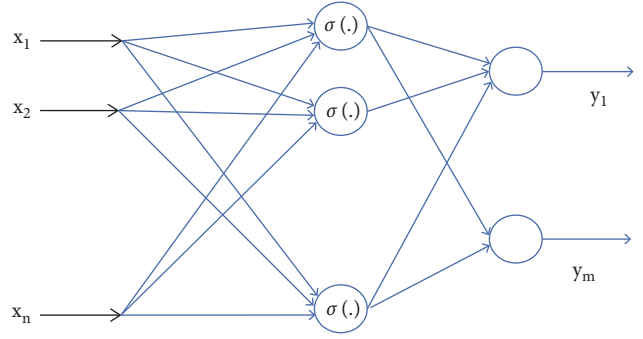
FIGURE 1: Feedforward neural network.

where $\sigma(x) = [\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_n)]^T$ is the activation function vector, $x \in R^n$.

The deviation weight value is used as the first column parameter of the weight matrix. In order to accommodate the deviation weight value, the vectors y and $\sigma(.)$ need to be expanded accordingly.

Property 1. Set the smoothing function from R^n to R^m to be h(y) and $U_y \subseteq R^n$. Then, under the condition of $y \in U_y$ and $\varepsilon > 0$, there are N hidden layer neurons and weight matrices W and V, and the relationship is $h(y) = W^T \sigma(VTy) + \varepsilon$. Then, the estimation of the smoothing function $h(y)$ can be expressed as $\hat{h}(y) = \hat{W}^T \sigma(\hat{V}^T y)$, and $\hat{W}$ and $\hat{V}$ are the estimation matrices of ideal neuron weights obtained by the online weight adjustment algorithm.

3.3. Engineering Mechanical Arm Neural Network Control

3.3.1. Neural Network Control. NNC (neural network control), also known as artificial neural network control system, neural network-based control, or neural control for short, refers to the use of neural network to realize the modeling, control, and optimization calculation of complex nonlinear objects. Such a control system is called neural network control system. The main system form of neural network control system is negative feedback regulation. The basic structure of the system is divided into open loop and closed loop. Its general structure is shown in Figure 2. The controller, identifier, and feedback link in the figure can be composed of neural network. Neural network uses binary logic, which is a digital network. Neural control is digital control, which uses digital quantity to control the controlled object. The neural network control system is a kind of digital control system, which has the general characteristics of the digital control system. Like the general microcomputer control system, it also has two parts: hardware and software. Neural network control has fundamentally changed the design idea of the traditional control system and has become a control method without mathematical model of controlled image [20].

In the control system, the neural network functions either as a controller or as an identifier. Neural controller has intelligent performance, so the neural network control system, as an intelligent control system, is a system with learning ability, also known as learning control system. The
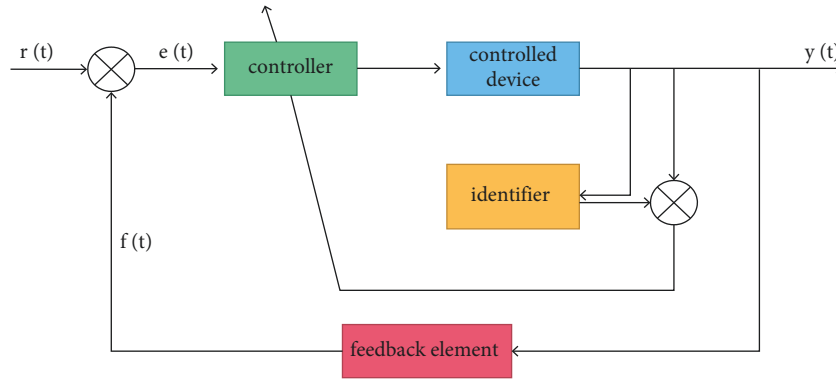
FIGURE 2: General structure of neural network control system R (T).
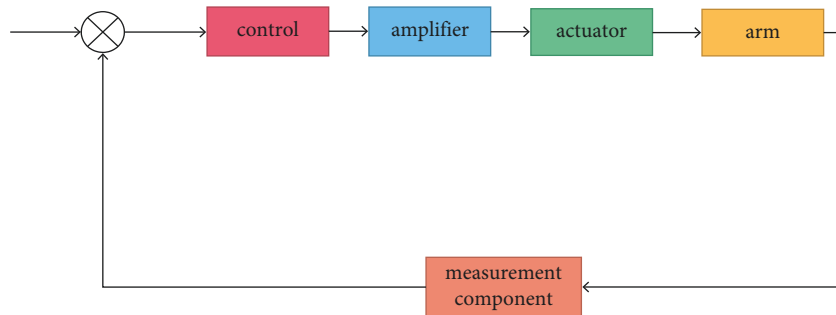
FIGURE 3: Control structure of mechanical arm.

learning process is a process of training and memorizing the training results. Compared with classical controller and modern controller, neural controller has obvious advantages and disadvantages. The biggest advantage is that the design of neural controller has nothing to do with the mathematical model of the controlled object, which is the fundamental reason why neural network can stand in automatic control. The disadvantage is that the neural network needs to carry out learning and training online or offline and use the training results to design the system. This kind of training largely depends on the accuracy of training samples, and the selection of training samples still has human factors. Neural identifier is used to identify the nonlinearity and uncertainty of the controlled object.

3.3.2. Neural Network Control Method of Engineering Mechanical Arm. The basic control of the mechanical arm includes position control, torque control, and force/position hybrid control. Most engineering mechanical arms adopt the control structure based on joint space. Firstly, the joint variables are obtained through inverse kinematics, and then the servo control is realized with the expectation of joint displacement, velocity, and acceleration. The control structure is shown in Figure 3.

In this paper, the research of engineering mechanical arm control starts from position control. Position control mainly refers to the method of the control system when the engineering mechanical arm tracks and controls the trajectory. These include control systems based mainly on switching variables and position-based control systems. The former is studied in joint space

and the latter is in operation space. Among them, the control system with shutdown variable is widely used.

In the control system based on joint variables, $q_d = [q_{d1}, q_{d2}, \ldots, q_{dn}]$ is the expected joint position vector, $\dot{q}_d$ and $\ddot{q}_d$ are the expected joint velocity and acceleration, u_1 and u_2 are control quantities, and $f = [f_1, f_2, \ldots, f_n]$ is the driving torque vector of the joint. The function of the control system is to find the appropriate control quantity to determine the driving torque required by the driving joint, so as to achieve the purpose of tracking the desired trajectory [21]. The above control process is based on the dynamics of the mechanical arm, so the control problem turns to the dynamics control of the mechanical arm. The traditional dynamics control structure is shown in Figure 4.

In Figure 5, q, $\dot{q}$ represent the joint position and speed of the engineering mechanical arm. As can be seen from the figure, the engineering mechanical arm controller includes a transmitter controller and a feedback controller, the most important part of which is the design of the transmitter controller. How the controller is designed to support the operation of the control system depends largely on the dynamic design of the engineering mechanical arm. If the model is clearly known, the traditional methods such as moment calculation method and accurate linearization method can meet the requirements. However, when the model is unknown, the control system actually becomes the traditional PID control. Obviously, such a control system is difficult to obtain good control performance. Using certain methods to eliminate or reduce the influence of the uncertainty of the engineering mechanical arm dynamic model on the control system, so as to improve the performance of
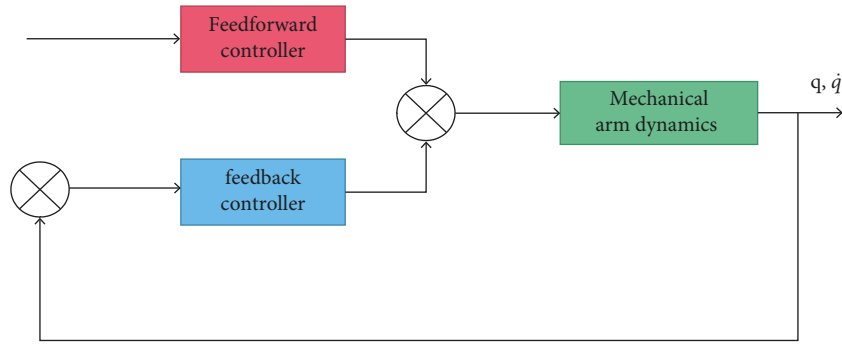
FIGURE 4: Dynamic control structure of mechanical arm.
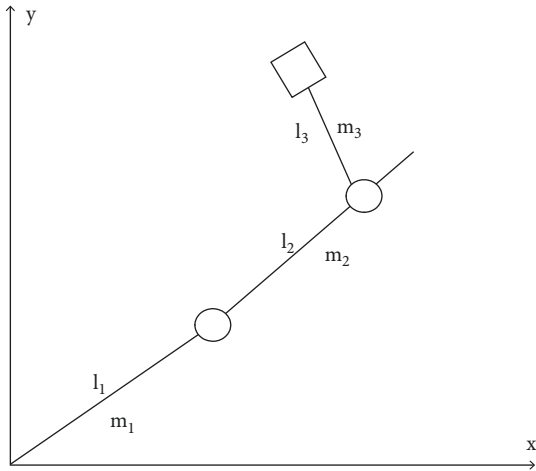
FIGURE 5: Schematic diagram of mechanical arm model.

—— Expect the trajectory
—— Neural network tracks the trajectories
—— PID tracks the trajectory

FIGURE 6: Movement track of engineering mechanical arm 3 (without external interference).

the control system, has become a problem that researchers need to solve. Thus, the adaptive method is effectively applied to it. Adaptive control can effectively overcome the structural uncertainty, compensate the influence of nonlinear and uncertain factors on the system, and can significantly improve the performance of the engineering mechanical arm. However, adaptive control is not effective for nonstructural engineering manipulator. In order to solve this problem, researchers have made a lot of exploration and opened up a new idea of control with the help of neural network. Neural network can become a powerful tool for engineering mechanical arm control, which is determined by its strong nonlinear mapping ability, self-learning ability, and adaptive ability [22, 23].

4. Result Discussion

The established schematic diagram of the engineering mechanical arm model is simulated and verified, as shown in Figure 5.

The tracking error of the engineering mechanical arm movement trajectory was modeled in the Matlab/Simulink environment to test the tracking effect of the mechanical arm motion trajectory controlled by the adaptive neural network. The simulation results of motion track tracking error of engineering mechanical arm 3 are obtained. The simulation

results of mechanical arm 1 and mechanical arm 2 are similar to those of mechanical arm 3, so simulation verification is not performed here. The simulation parameters are set as follows: the expected trajectory of mechanical arm 3 is $\theta_3 = 0.4 \cos(2\pi t)$, the initial condition is $\theta(0) = [000]^T$, the control parameter $k = \text{diag}(40, 40, 40)$, the interference parameter $\tau' = 20 \cos(\pi t)$, and the mechanical arm link parameters are $L1 = 0.62$ M, $L2 = 0.41$ m, $L3 = 0.34$ M, $M1 = 3.5$, $M2 = 2.5$ kg, $m3 = 2.0$ kg, $g = 9.82$ m/s^2, and $t = 2$ S. Under the condition of no external interference, the simulation results of motion trajectory tracking of engineering mechanical arm 3 are shown in Figure 6, and the simulation results of input torque of engineering mechanical arm 3 are shown in Figure 7. In the case of external interference, the trajectory tracking simulation results of engineering mechanical arm 3 are shown in Figure 8. The simulation results of input torque of engineering mechanical arm 3 are shown in Figure 9.

It can be seen from Figures 6 and 7 that under the condition of no external interference, both adaptive neural
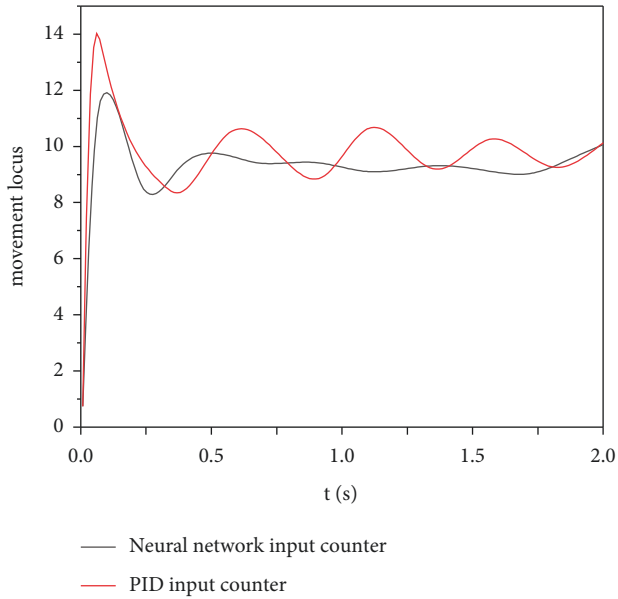
Figure 7: Torque control of engineering mechanical arm 3 (no external interference).
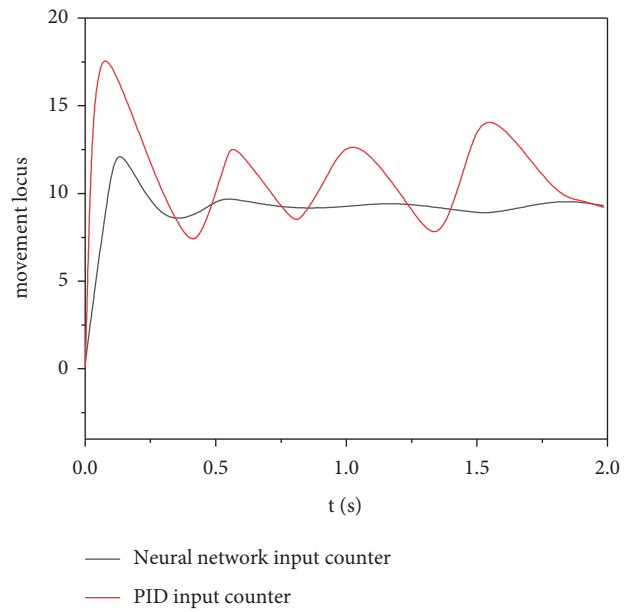
Figure 9: Torque control of engineering mechanical arm 3 (with external interference).
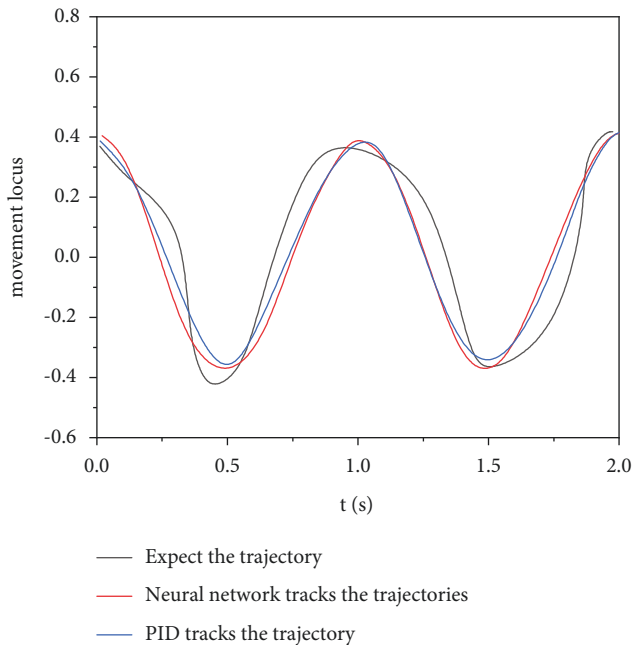
Figure 8: Movement track of engineering mechanical arm 3 (with external interference).

network control and PID control can effectively track the desired trajectory of the engineering mechanical arm. Using PID control, the input torque fluctuates greatly and the jitter phenomenon is serious. Therefore, the adaptive neural network control method is better. It can be seen from Figures 8 and 9 that under the condition of external interference, the adaptive neural network control can better track the desired trajectory of the engineering mechanical arm, and the fluctuation range of input torque is small. However, the PID control has a very large tracking error in the expected trajectory path of the engineering

mechanical arm, the range of input torque fluctuations is up to 20, and the jitter phenomenon is more serious. Adaptive neural network monitoring can be used to control the trajectory of engineering manipulator.

5. Conclusion

The modeling and control of a nonlinear dynamic system based on neural network is a frontier subject of nonlinear discipline, which is of great significance in many fields, especially for the multi-variable and strong disaster uncertain nonlinear object such as mechanical arm. This article examines the control of the neural network to control the trajectory of the engineering mechanical arm. The neural network controller can effectively control the pathway of the end effector and control the subtask of the program. Given the dynamic equations of the engineering mechanical arm, the dynamic properties of the engineering mechanical arm were studied using a positive feedback network. The adaptive neural network management system was developed, and the stability and integrity of the closed-loop system were proved by Lyapunov's function. We created a schematic diagram of the engineering mechanical arm model and simulated the dynamic parameters of the engineering mechanical arm using the Matlab/Simulink program. In addition, the PID control system simulation results are compared and analyzed. The simulation results show that in the case of external interference, the adaptive neural network control method can not only accurately realize the trajectory tracking task of the mechanical arm but also effectively weaken the jitter of the engineering mechanical arm.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

Acknowledgments

References

[1] K. Monkova, M. Vasina, M. Zaludek, P. P. Monka, and J. Tkac, "Mechanical vibration damping and compression properties of a lattice structure," *Materials*, vol. 14, no. 6, p. 1502, 2021.

[2] Y. Yahagi, S. Fukushima, S. Sakaguchi, and T. Naemura, "Suppression of floating image degradation using a mechanical vibration of a dihedral corner reflector array," *Optics Express*, vol. 28, no. 22, p. 33145, 2020.

[3] Z. Zhao, K. Wei, W. Ding, F. Cheng, and P. Wang, "Evaluation method of the vibration reduction effect considering the real load- and frequency-dependent stiffness of slab-track mats," *Materials*, vol. 14, no. 2, p. 452, 2021.

[4] M. Khazaee, M. A. Mirsalim, S. Minaei, A. Banakar, B. Ghobadian et al., "Remaining useful life (RUL) prediction of internal combustion engine timing belt based on vibration signals and artificial neural network," *Neural Computing & Applications*, vol. 33, no. 13, pp. 7785–7801, 2021.

[5] H. S. Yang, C. Y. Dong, X. C. Qin, and Y. H. Wu, "Vibration and buckling analyses of fgm plates with multiple internal defects using xiga-pht and fcm under thermal and mechanical loads," *Applied Mathematical Modelling*, vol. 78, pp. 433–481, 2020.

[6] M. Pöschl, M. Vašina, P. Zádrapa, D. Měřínská, and M. Žaludek, "Study of carbon black types in sbr rubber: mechanical and vibration damping properties," *Materials*, vol. 13, no. 10, p. 2394, 2020.

[7] E. Fuhrer, M. Jouda, C. O. Klein, M. Wilhelm, and J. G. Korvink, "Gradient-induced mechanical vibration of neural interfaces during mri," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 915–923, 2020.

[8] K. C. Schvartz-Leyzac, Y. Raphael, B. E. Pfingst et al., "How electrically evoked compound action potentials in chronically implanted Guinea pigs relate to auditory nerve health and electrode impedance," *Journal of the Acoustical Society of America*, vol. 148, no. 6, pp. 3900–3912, 2020.

[9] L. Ghasemi-Mobarakeh, M. P. Prabhakaran, M. Morshed, M. H. Nasr-Esfahani, and S. Ramakrishna, "Electrospun poly(ε-caprolactone)/gelatin nanofibrous scaffolds for nerve tissue engineering," *Biomaterials*, vol. 29, no. 34, pp. 4532–4539, 2008.

[10] L. H. Carney, M. J. Mcduffy, and I. Shekhter, "Frequency glides in the impulse responses of auditory-nerve fibers," *Journal of the Acoustical Society of America*, vol. 105, no. 4, pp. 2384–2391, 1999.

[11] S. H. Hsu, S. H. Chan, C. M. Chiang, C. Chi-Chang Chen, and C. F. Jiang, "Peripheral nerve regeneration using a microporous polylactic acid asymmetric conduit in a rabbit long-gap sciatic nerve transection model," *Biomaterials*, vol. 32, no. 15, pp. 3764–3775, 2011.

[12] P. J. Abbas and M. B. Sachs, "Two-tone suppression in auditory-nerve fibers: extension of a stimulus-response relationship," *Journal of the Acoustical Society of America*, vol. 59, no. 1, pp. 112–122, 1976.

[13] L. Tai and J. J. Guinan, "Lin, t. & guinan, j.j. auditory-nerve-fiber responses to high-level clicks: interference patterns indicate that excitation is due to the combination of multiple drives," *Journal of the Acoustical Society of america*, vol. 107, no. 5 Pt 1, pp. 2615–2630, 2000.

[14] J. Patocka, D. Jun, J. Bajgar, and K. Kuca, "Prophylaxis against nerve agent intoxications," *Defence Science Journal*, vol. 56, no. 5, pp. 775–784, 2006.

[15] G. H. Liu, H. Hu, P. H. Zhang, and W. Z. Wang, "Radial compressive properties of the biodegradable braided regeneration tubes for peripheral nerve repair," *Journal of Industrial Textiles*, vol. 36, no. 1, pp. 35–46, 2006.

[16] N. C. Blacker, P. H. Findlay, and D. C. Sherrington, "Synthesis of cu-ii-complexed polymers and use as catalysts in the hydrolytic decontamination of sarin nerve agent," *Polymers for Advanced Technologies*, vol. 12, no. 3-4, pp. 183–196, 2001.

[17] Y. Fu, M. Tang, X. D. Sun, and X. Xu, "Thickness of retinal nerve fiber layer and central fovea of macula in children with unilateral amblyopia," *Journal of Shanghai Jiaotong University*, vol. 32, no. 2, pp. 235–237, 2012.

[18] Q. Liu, J. Huang, H. Shao, L. Song, and Y. Zhang, "Dual-factor loaded functional silk fibroin scaffolds for peripheral nerve regeneration with the aid of neovascularization," *RSC Advances*, vol. 6, no. 9, pp. 7683–7691, 2016.

[19] D. J. Knauer, K. M. Scaparro, and D. D. Cunningham, "The gamma subunit of 7 S nerve growth factor binds to cells via complexes formed with two cell-secreted nexins," *Journal of Biological Chemistry*, vol. 257, no. 24, pp. 15098–15104, 1982.

[20] D. I. Stephanova, "Conduction along myelinated and demyelinated nerve fibres during the recovery cycle: model investigations," *Biological Cybernetics*, vol. 62, no. 1, pp. 83–87, 1989.

[21] T. Guan, G. Nie, J. Wang, M. Yang, K. Zhu, and Y. Wang, "Near-infrared laser stimulation of the auditory nerve in Guinea pigs," *Journal of the Optical Society of Korea*, vol. 20, no. 2, pp. 269–275, 2016.

[22] Y. Bian, X. Liang, and Z. Gao, "Vibration reduction for a flexible arm using magnetorheological elastomer vibration absorber," *Shock and Vibration*, vol. 2018, pp. 1–13, Article ID 9723538, 2018.

[23] D. Grzybek, H. Pamula, J. Huebner et al., "Piezoelectric particulate composite for energy harvesting from mechanical vibration," *Materials*, vol. 13, no. 21, p. 4925, 2020.

Retraction

Retracted: Big Data Value Calculation Method Based on Particle Swarm Optimization Algorithm

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] W. Ma and X. Hou, "Big Data Value Calculation Method Based on Particle Swarm Optimization Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5356164, 8 pages, 2022.

Research Article

Big Data Value Calculation Method Based on Particle Swarm Optimization Algorithm

Wensheng Ma[1] and Xilin Hou [iD][2]

[1]*School of Electronic and Information Engineering, Liaoning University of Science and Technology, Anshan 114051, Liaoning, China*
[2]*School of Business Administration, Liaoning University of Science and Technology, Anshan 114051, Liaoning, China*

Correspondence should be addressed to Xilin Hou; 191081100015@stu.ustl.edu.cn

SI is a relatively recent technology that was inspired by observations of natural social insects and artificial systems. This system comprises multiple individual agents who rely on collective behavior in decentralized and self-organized networks. One of the biggest difficulties for existing computer techniques is learning from such large datasets, which is addressed utilizing big data. Big data-based categorization refers to the challenge of determining which set of classifications a new discovery belongs to. This classification is based on a training set of data that comprises observations that have been assigned to a certain category. In this paper, CIN-big data value calculation based on particle swarm optimization (BD-PSO) algorithm is proposed by operating in local optima and to improve the operating efficiency. The convergence speed of the particle swarm optimization (PSO), which operates in the local optima, is improved by big data-based particle swarm optimization (BD-PSO). It improves computing efficiency by improving the method, resulting in a reduction in calculation time. The performance of the BD-PSO is tested on four benchmark dataset, which is taken from the UCI. The datasets used for evaluation are wine, iris, blood transfusion, and zoo. SVM and CG-CNB are the two existing methods used for the comparison of BD-PSO. It achieves 92% of accuracy, 92% of precision, 92% of recall, and 1.34 of F1 measure, and time taken for execution is 149 ms, which in turn outperforms the existing approaches. It achieves robust solutions and identifies appropriate intelligent technique related to the optimization problem.

1. Introduction

In this day and age, the development of high-throughput technologies has resulted in exponential increase in harvested information [1]. This exponential growth termed as "big data" (BD) is in terms of both dimensionality and sample size. Nowadays, efficient and effective management of these big data is increasingly challenging. Traditional management techniques of these become impractical [2]. Therefore, data mining (DM), machine learning (ML), and metaheuristic techniques are developed to automatically discover knowledge and recognize patterns from these big data [3, 4].

The categorization of big data (BD) is an essential procedure that aids in the effective study of enormous

datasets [5]. For effective BD classification, highly parallelized learning algorithms must be designed. Many relevant data features, such as a high-dimensional dataset, a large number of data kinds (classes), high-speed data processing, and unstructured data, make up the complexity parameter of big data [6]. Machine learning approaches will be used to solve the complexity parameter, and the certain difficulties it causes are hard to handle. However, upgrading present learning algorithms to deal with massive data categorization challenges and needs remains a difficulty [7]. The process of big data is given in Figure 1.

Evolutionary computation (EC) approaches have been applied to scheduling difficulties, resulting in the
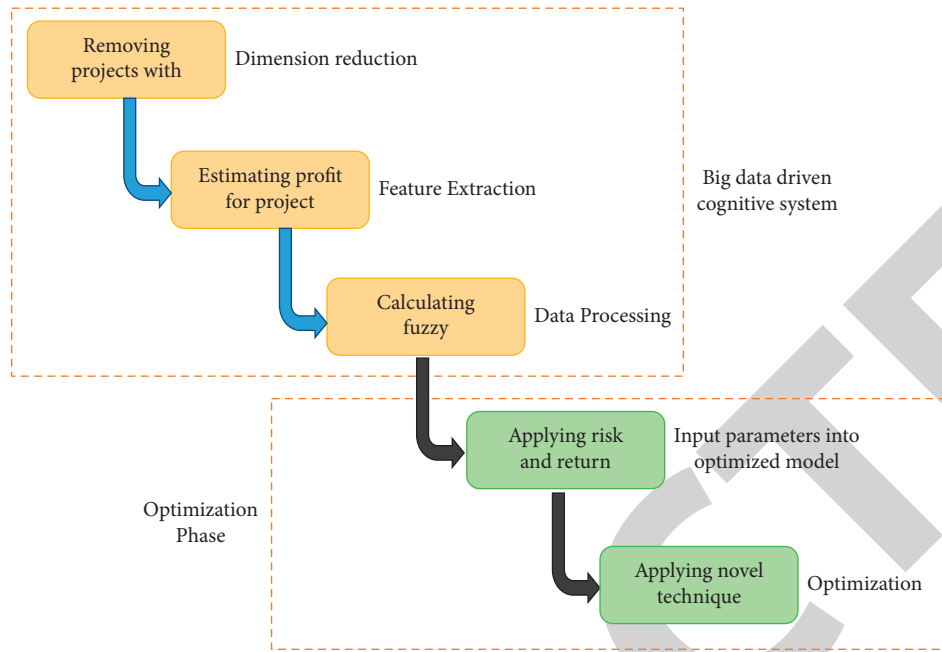
Figure 1: Methodology of big data.

evolutionary scheduling (ES) study field. EC is a rapidly expanding artificial intelligence (AI) study topic [8, 9]. Natural selection and genetic inheritance are examples of EC approaches that draw concepts and inspiration from natural evolution/adaptation. Evolutionary algorithm (EA) and structural inference (SI) are the two basic categories of EC. The SI is a new field of research for EC [10, 11]. It is a novel computational and behavioral paradigm for addressing scheduling issues that was identified via simplified social behaviors of insects and other animals. It is inspired by the collective intelligence of swarms of biological populations [12].

Most of the optimization algorithms suffer due to the exploitation and exploration problem, since the identification of target entirely depends on the initial solution of the optimization algorithm [13, 14]. The same issue appears in PSO and some other optimization search algorithm also. These two optimization algorithms are more advantageous than other existing optimization algorithms. In PSO, there will be no adaptive variation or random solution so that the generation of fresh solution takes place around the initial solution [15].

The exploration issue entails condensing a large number of implausible answers into a single group and selecting the best among them. The exploitation of the problem, on the other hand, is concerned with finding the best solution among the many possibilities [16]. PSO algorithm also has some advantages, which can generate optimized results. One of the significant advantages of PSO is that the new solution generated is based on the local best and global best. This will deliver the new solution by considering the best solution of current and whole iterations, so that fresh solution can travel to the target in a smoother way [17].

This paper is motivated to calculate CIN-big data value based on particle swarm optimization (BD-PSO) algorithm operating in local optima and to improve the operating efficiency.

The remainder of the research article is systemized as follows: recent works in big data classification are given in Section 2, the proposed methodology is discussed in Section 3, the outcome is compared and contrasted in Section 4, and the article is concluded in Section 5.

2. Related Works

The supervised machine learning-based classification approaches to train the data for effective classification problems. In the quantum computer, a binary classifier approach for optimization issues is introduced; that is, support vector machine and the complexity of the computation are reduced. The matrix inversion approach is used in training the matrix, and exponentiation of nonsparse matrix is the core concept of the quantum big data approach [18]. The intrusion in the network traffic is minimized by the continuous collection and processing of collected data. The continuous collection of data results in the growth huge volume of data, and the machine learning approach is used in processing and formulating the significant inference from the data [19].

The processing and maintaining of big data necessitate the robust technique where the shortcomings of the traditional approaches are rectified by the training and learning approaches [20]. Big data is an eminent research and application field, whereby the extraction of significant insight with high scalability is a complicated process. The issue of scalability is rectified by MapReduce, and it utilizes the divide and conquer method. The scalability of the big

data is rectified by the MapReduce framework [21]. The ensembles of particle swarm optimization (PSO) and support vector machine (SVM) are considered to classify the big data, and significant insights are acquired from the classified data [22].

In the problem of classification, the feature selection approach is incorporated with the optimization approach to attain an effective solution. The cat swarm optimization (CSO) is developed from the food searching behavior of the cat, and it is modified to classify the big data [23]. Big data has shown its progression in diverse industry and application domains, whereas the growth of data necessitated a strong approach to process those data. Cuckoo-grey wolf-based correlative Naive Bayes (CG-CNB) classifier is framed by altering the CNB classifier with a developed optimization approach that is cuckoo-grey wolf-based optimization (CGWO).

The CGCNB-MRM method executes the process of classification for every data sample based on the posterior probability of the data and the probability index table [24]. The scale-free particle swarm optimization (SF-PSO) is developed for the selection of features in the high-dimensional datasets. The multi-class support vector machine (MC-SVM) is incorporated as a machine learning classifier and acquired the best result. The big data classification approaches have numerous drawbacks that are rectified with the incorporation of optimization and the deep learning-based approach.

The extreme learning machine (ELM) and particle swarm optimization (PSO) were integrated in [25] to select features and to determine the hidden node's count. The classification of sleep stages was used for predicting the proportion of sleep stages. Support vector machine (SVM) is used for comparing ELM methods that are lesser than the ELM and PSO integration. In [26], PSO algorithm was presented for performing a global search for the optimal weights/biases of the ANN method selected. This method is represented as PSO-ANN. A performance metric's variety was utilized for assessing the training procedure quality, and also the performance of model in the testing dataset. The results exposed that the representations established and GCV was determined accurately and rapidly.

3. Big Data-Based Particle Swarm Optimization Algorithm

The appliance of PSO is developed from the complex social behavior which is exposed by the natural species. For a D-dimensional search space, the position of the ith particle is denoted as $X_i = (x_{i1}, x_{i2}, \ldots \ldots x_{iD})$. Each particle upholds a memory of its previous best position $P_i = (p_{i1}, P_{i2}, \ldots, \ldots p_{iD})$ and a velocity $V_i = (v_{i1}, v_{i2}, \ldots v_{iD})$ along each dimension. The P vector of the particle with the best fitness in the immediate neighborhood is nominated as g at each iteration. The current particle's P vector is merged to change the velocity along each dimension, and the particle's new location is computed using the adjusted velocity. The conventional PSO Algorithm 1 pseudo-code is shown below.

PSO is a population-based stochastic search algorithm inspired by the social behavior of a flock of birds. This method has a population of particles, each of which represents a possible solution to the issue. A swarm in the context of PSO indicates to a group of potential results to an optimization issue, and every outcome can be indicated as a particle. The PSO's main purpose is to locate the particle location that yields the best assessment of a fitness function. Every particle indicates a location in N_d-dimensional space, and it is flown across this multi-dimensional search space, altering its position in relation to both other particles until the optimal position is identified. Each particle I is responsible for maintaining the subsequent data.

xc_i: Particle's exact position.

ve_i: Particle's exact velocity.

yp_i: Particle's best position (personal).

By utilizing the above notation, a particle's incidence is altered by

$$ve_i, k(t+1) = wve_i + k(t)c_1 r_{1.k}(t)\left(p_{i,k}(t) - xc_{i,k}(t)\right)$$
$$+ c_2 r_{2.k}(t)\left(p_{k,g}(t) - xc_{i,k}(t)\right) \tag{1}$$
$$xc_i(t+1) = xc_i(t) + ve_i(t+1),$$

where the inertia is indicated by ω, acceleration constants are indicated as c_1 and c_2, $r1.(t)$, $r2.kt \sim U(0.1)$, and $k = 1, \ldots, Nd$.

The velocity is thus computed based on the three influences which are described below.

(i) A fraction of the former velocity.

(ii) The cognitive module, which is considered as a function of the distance of the particle from its private best position.

(iii) The social component, which is determined by the particle's distance from the best particle identified thus far.

The pbest$_i$ value is presented as the best previously visited position of the ith particle. It is signified as $pi = (pi1, pi2, \ldots, pi D)$. The $g$best is the global best position of the all individual $p$best$i$ values. It is represented as the $g = (g1, g2, \ldots, g D)$. The position of the ith particle is denoted by $(x_{il}, x_{i2}, \ldots, x_{i d})$, $x \in (x_{\min}, x_{\max}).D$, and its velocity is indicated by $vi = (vi1, vi2, \ldots vi D)$ $v \in (v\min, v\max).D$.

$$v_{\text{id}}^{\text{new}} = w * v_{\text{id}}^{\text{old}} + c_1 * r_1 * \left(pbest_{\text{id}} - x_{\text{id}}^{\text{old}}\right)$$
$$+ c_2 * r_2 * \left(pbest_{\text{id}} - x_{\text{id}}^{\text{old}}\right) \tag{2}$$
$$x_{\text{id}}^{\text{new}} = x_{\text{id}}^{\text{old}} + v_{\text{id}}^{\text{new}},$$

where r_1 and r_2 represent the random count among (0, 1); c_1 and c_2 control how far a particle will move in one's generation; and $v_{\text{id}}^{\text{new}}$ and $v_{\text{id}}^{\text{old}}$ denote the old and new particle, respectively. The existing particle position is $x_{\text{id}}^{\text{old}}$, while the revised particle position is $x_{\text{id}}^{\text{new}}$. The inertia weight w

```
(1) PSO procedure
(2) Parameter and population initialization
(3) For every individual particle
(4) Investigate objective function
(5) Update local and global best
(6) For each particle do
(7) Update velocity and position
(8) Investigate objective function
(9) Update local and global best
(10) While
(11) End procedure
```

ALGORITHM 1: Procedure for PSO.

regulates the effect of a particle's prior velocity on its current velocity. w is intended to exchange V_{max} and alter the optimization process's influence of prior particle velocities. An acceptable compromise between exploration and exploitation is critical for high-performance issues. How to optimally stabilize the swarm's search skills is one of the most important issues in PSO because maintaining a good mix search during the whole run is complex to PSO's performance. Throughout the search process, the inertia weight decreases linearly from 0.9 to 0.4, and the specific equation may be expressed as follows:

$$W_{LDW} = (w_{max} - w_{min}) * \frac{\text{iteration}_{max} - \text{iteration}_i}{\text{iteration}_{max}} + w_{min}, \tag{3}$$

where w_{max} is 0.9, w_{min} is 0.4, and Iteration$_{max}$ is the maximal count of permitted iterations.

Contrarily, these computation tasks consume more time. Therefore, in between, optimal scheduling of tasks to the nodes is performed by utilizing the same optimization algorithms. This helps in determining allocation of the tasks efficiently to the corresponding nodes using a random approach, thus reducing the overall computation time. All these processes are employed using the MR programming model, for parallelization purpose. This is one of the most popular distributed processing systems implemented within the Hadoop environment. It provides a design pattern that advises algorithms to be represented in two fundamental functions, known as "Map" and "Reduce" to ease enormous simultaneous processing of large datasets. "Map" is used for per-record calculation in the first phase, which means that the input data are handled by this function to provide some intermediate results. The intermediate outputs are then passed into a second phase known as the "Reduce" function, which combines the output from the Map phase and applies a specific function to get the final results.

4. Result and Discussion

In this research work, big data-based calculation is done by using a particle swarm optimization (PSO). This approach is implemented and tested using MATLAB. The performance of this algorithm is evaluated using different datasets and compared with the outcome acquired by utilizing various existing optimization algorithms. The datasets used for evaluation are wine, iris, blood transfusion, and zoo. The performance of the BD-PSO is tested on four benchmark datasets, which are taken from the UCI. The number of iterations taken to attain the best fitness value is 10. In the wine dataset, 178 samples are formed into three subclasses by applying the BD-PSO clustering algorithm.

4.1. True Positive. The proportion of correct forecasts in positive class predictions is known as the true-positive rate (TPR). The true-positive rate is given as

$$TP = \frac{TP}{TP + FN}. \tag{4}$$

4.2. True Negative. A true-negative test outcome in a study of diagnostic test accuracy indicates that the test being assessed accurately showed that a participant did not have the particular goals, according to the reference standardized score where individual did not have the negative condition. The true-negative rate is given as

$$TN = \frac{TN}{TN + FP}. \tag{5}$$

4.3. False Positive. The false-positive rate (FPR) is a statistic for assessing the accuracy of a test, being it an inquiry, a machine learning approach, or something else. In technical terms, the FPR refers to the chance of mistakenly rejecting the null hypothesis. The false-positive rate is calculated as follows:

$$FP = \frac{FP}{TN + FP}. \tag{6}$$

4.4. False Negative. A "false alarm" is a term for a false positive. When you suggest something is untrue when it is actually the truth, you are using a false negative. The false-negative rate is given as

Table 1: Comparison of accuracy.

File name	SVM	CG-CNB	BD-PSO
Wine	73.18	81.87	92.03
Iris	74.61	82.98	92.65
Blood transfusion	74.96	81.26	89.25
Zoo	76.36	82.86	91.94

$$FN = \frac{FN}{TP + FN}. \tag{7}$$

4.5. Accuracy. The classification accuracy of the instance is estimated by dividing the count of appropriate negative instance identifications by the total count of instance. The competency of the classification model is determined by the accuracy value. The accuracy is measured using true-positive (TP) and true-negative (TN) values generated from instance-based classes. Comparison of accuracy is given in Table 1 and Figure 2. The most accurate classification method is known as an effective classification algorithm. An estimate of the accuracy value is as follows:

Figure 2: Comparison of accuracy.

$$Accuracy = \frac{True\ Positive\,(TP) + True\ Negative\,(TN)}{True\ Positive\,(TP) + True\ Negative\,(TN) + False\ Positive\,(FP) + False\ Negative\,(FN)}. \tag{8}$$

4.6. Precision. The quantitative rate with positive results, also known as precision, reflects the reliability of the prediction and the relevance of the feature found. The frequency of arbitrary mistakes is expressed as precision, which is expressed using statistical variables. Accuracy and precision are the same concepts for received feature values. Typically, binary or decimal digits are used to represent the accuracy of a data. True-positive (TP) and false-positive (FP) rates are used to calculate it. The fraction of positive values in the whole fake profile determines the precision value. The count of genuine positive attributes is the accuracy count for a certain issue in the categorization process (i.e., the count of the item relevantly labeled as positive classes of instances). The application's high precision shows a resulting value that achieves more desired data than the incorrect data. Comparison of precision is given in Table 2 and Figure 3. It is equated as

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}. \tag{9}$$

4.7. Recall. The associated fake profiles among the substantially retrieved occurrences make up the rate of recall. The estimated measure of recall is effective in predicting rate, and recall is the count of associated events. In the fake profiles, recall is calculated as the count of accurately
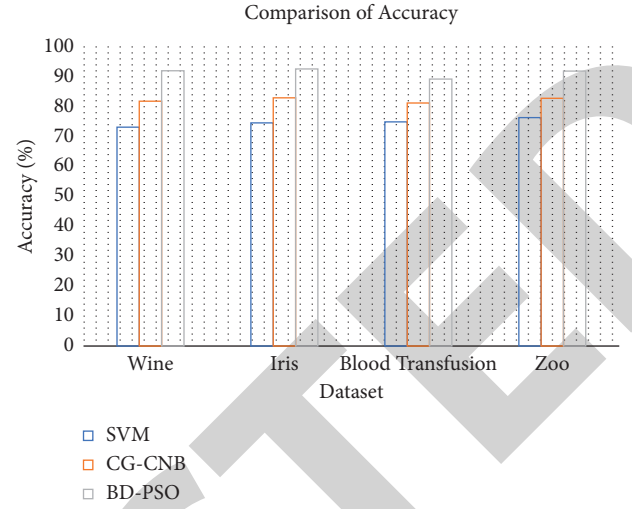
detected values divided by the count of TP and FP data. The difference in the negative instance recognition at TP and FN rates is used to assess precision. Comparison of recall is given in Table 3 and Figure 4. It is calculated as

$$Recall = \frac{TP}{TP + FN}. \tag{10}$$

4.8. F-Measure. The accuracy of the examination of the categorization problem is indicated by the *F*-measure or *F*-score. The method achieves the optimum *F*-measure value by achieving the highest accuracy and recall value. The *F*-measure value improves the extraction of essential information from characteristics and provides an accurate representation of the computation performance. Comparison of *F*-measure is given in Table 4 and Figure 5. It is computed as

$$F - measure = \frac{2\Delta\ Precision\ \Delta\ Recall}{Precision + Recall}. \tag{11}$$

4.9. Execution Time. The time taken to complete the classification process is determined as execution time, and the algorithm with minimal execution time is the effective algorithm. The values of proposed and existing approach are given in Table 5 that is illustrated in Figure 6.

The numerical outcome of the proposed approach is given in above tables and figures. The results are compared for different files, and the outcome of the proposed approach

Table 2: Comparison of precision.

File name	SVM	CG-CNB	BD-PSO
Wine	74	76	88
Iris	70	79	89
Blood transfusion	81	77	92
Zoo	76	80	89

Table 4: Comparison of *F*-measure.

File name	SVM	CG-CNB	BD-PSO
Wine	2.174089	2.531062	1.261859
Iris	2.15625	2.338357	1.265407
Blood transfusion	2.455046	2.418357	1.341966
Zoo	2.275248	2.238377	1.32867

Figure 3: Comparison of precision.
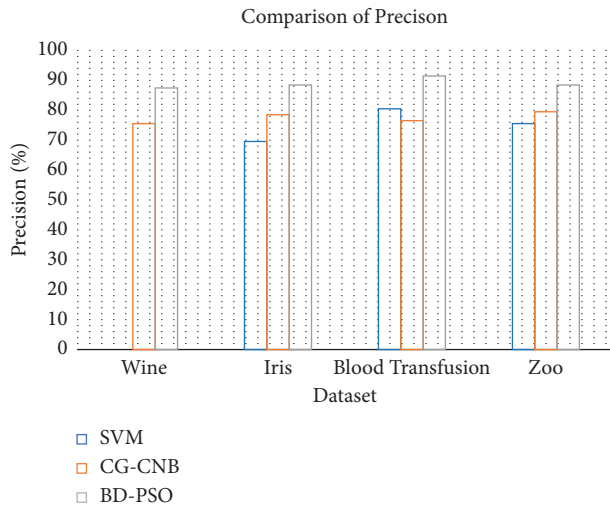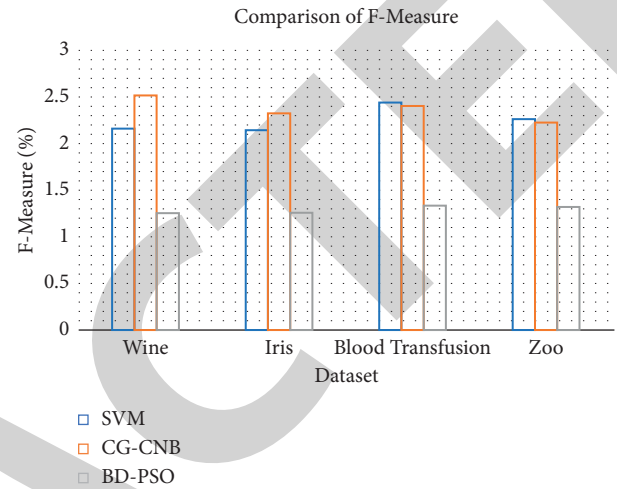
Figure 5: Comparison of *F*-measure.

Table 3: Comparison of recall.

File name	SVM	CG-CNB	BD-PSO
Wine	72	77	91
Iris	71	78	89
Blood transfusion	81	81	92
Zoo	75	80	89

Table 5: Comparison of *F*-measure.

File name	SVM	CG-CNB	BD-PSO
Wine	228	199	161
Iris	219	187	158
Blood transfusion	186	196	149
Zoo	223	179	163

Figure 4: Comparison of recall.
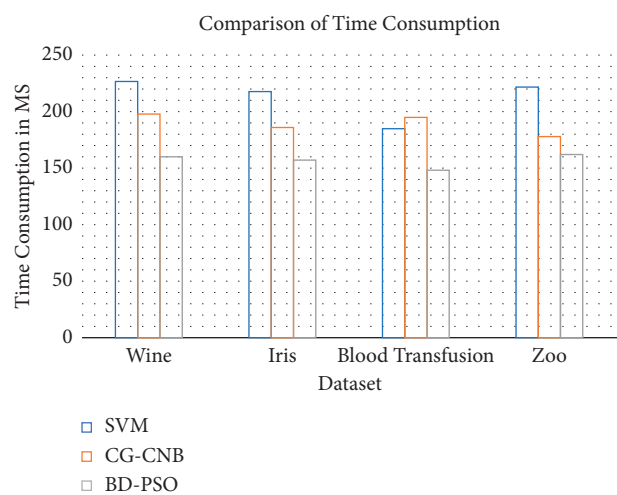
Figure 6: Comparison of execution time.

outperforms the existing approaches. The performance of the BD-PSO is tested on four benchmark datasets, which are taken from the UCI. The datasets used for evaluation are wine, iris, blood transfusion, and zoo. SVM and CG-CNB are the two existing methods used for the comparison of BD-PSO. It achieves 92% of accuracy, 92% of precision, 92% of

recall, and 1.34 of F1 measure, and time taken for execution is 149 ms, which in turn outperforms the existing approaches. However, the results obtained do not provide optimal solution.

5. Conclusion

Swarm intelligence (SI) is a relatively new technology derived from observations of natural social insects and artificial systems. In decentralized and self-organized systems, this system consists of several individual agents who rely on collective behavior. Learning from such big datasets is one of the major issues for the current computational algorithms, whereby the drawback is rectified using big data. The difficulty of identifying to which set of classes a new discovery belongs is known as big data-based categorization. This identification is based on a training set of data that encompasses interpretations with identified class membership. BD-PSO enhanced the speed of convergence in the PSO, which runs in the local optima. The performance of the BD-PSO is tested on four benchmark datasets, which are taken from the UCI. The datasets used for evaluation are wine, iris, blood transfusion, and zoo. SVM and CG-CNB are the two existing methods used for the comparison of BD-PSO. It achieves 92% of accuracy, 92% of precision, 92% of recall, and 1.34 of F1 measure, and time taken for execution is 149 ms, which in turn outperforms the existing approaches. It thereby increases the computational efficiency by optimizing the algorithm, thus reducing the computational time. It achieves robust solutions and identifies appropriate intelligent technique related to the optimization problem. In the future, the multi-objective big data based on hybrid optimization algorithm can be used for achieving optimal results.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest with any financial organizations regarding the material reported in this manuscript.

References

[1] A. K. Kushwaha, A. K. Kar, and Y. K. Dwivedi, "Applications of big data in emerging management disciplines: a literature review using text mining," *International Journal of Information Management Data Insights*, vol. 1, no. 2, Article ID 100017, 2021.

[2] C. Fan, D. Yan, F. Xiao, A. Li, J. An, and X. Kang, ". Advanced data analytics for enhancing building performances: from data-driven to big data-driven approaches," *Building Simulation*, vol. 14, no. 1, pp. 3–24, 2021.

[3] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, "A guide to machine learning for biologists," *Nature Reviews Molecular Cell Biology*, vol. 23, no. 1, pp. 40–55, 2022.

[4] B. Liu, M. Ding, S. Shaham, W. Rahayu, F. Farokhi, and Z. Lin, "When machine learning meets privacy: a survey and outlook," *ACM Computing Surveys*, vol. 54, no. 2, pp. 1–36, 2022.

[5] J. Wang, C. Xu, J. Zhang, and R. Zhong, "Big data analytics for intelligent manufacturing systems: a review," *Journal of Manufacturing Systems*, vol. 62, 2021.

[6] Y. Hajjaji, W. Boulila, I. R. Farah, I. Romdhani, and A. Hussain, "Big data and IoT-based applications in smart environments: a systematic review," *Computer Science Review*, vol. 39, Article ID 100318, 2021.

[7] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The role of ai, machine learning, and big data in digital twinning: a systematic literature review, challenges, and opportunities," *IEEE Access*, vol. 9, pp. 32030–32052, 2021.

[8] B. Doerr and F. Neumann, *Theory of Evolutionary Computation: Recent Developments in Discrete Optimization*, Springer, Berlin, Germany, 2019.

[9] Z. Wang and A. Sobey, "A comparative review between Genetic Algorithm use in composite optimisation and the state-of-the-art in evolutionary computation," *Composite Structures*, vol. 233, Article ID 111739, 2020.

[10] A. Telikani, A. H. Gandomi, and A. Shahbahrami, "A survey of evolutionary computation for association rule mining," *Information Sciences*, vol. 524, pp. 318–352, 2020.

[11] Z. H. Zhan, L. Shi, K. C. Tan, and J. Zhang, "A survey on evolutionary computation for complex continuous optimization," *Artificial Intelligence Review*, vol. 55, no. 1, pp. 59–110, 2022.

[12] A. Nayyar, S. Garg, D. Gupta, and A. Khanna, "Evolutionary computation: theory and algorithms," in *Advances in Swarm Intelligence for Optimizing Problems in Computer Science*, pp. 1–26, Chapman and Hall/CRC, London, UK, 2018.

[13] L. Abualigah, A. Diabat, S. Mirjalili, M. AbdElaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," *Computer Methods in Applied Mechanics and Engineering*, vol. 376, Article ID 113609, 2021.

[14] Z. Cui, Y. Cao, X. Cai, J. Cai, and J. Chen, "Optimal LEACH protocol with modified bat algorithm for big data sensing systems in Internet of Things," *Journal of Parallel and Distributed Computing*, vol. 132, pp. 217–229, 2019.

[15] I. M. El-Hasnony, S. I. Barakat, M. Elhoseny, and R. R. Mostafa, "Improved feature selection model for big data analytics," *IEEE Access*, vol. 8, pp. 66989–67004, 2020.

[16] D. Zhao and J. Liu, "Study on network security situation awareness based on particle swarm optimization algorithm," *Computers & Industrial Engineering*, vol. 125, pp. 764–775, 2018.

[17] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Physical Review Letters*, vol. 113, no. 13, Article ID 130503, 2014.

[18] S. Suthaharan, "Big data classification: problems and challenges in network intrusion prediction with machine learning," *ACM Sigmetrics-Performance Evaluation Review*, vol. 41, no. 4, pp. 70–73, 2014.

[19] S. Suthaharan, "Machine learning models and algorithms for big data classification," *Integrated Series in Information Systems*, vol. 36, pp. 1–12, 2016.

[20] A. Fernández, S. Del Río, N. V. Chawla, and F. Herrera, "An insight into imbalanced big data classification: outcomes and challenges," *Complex & Intelligent Systems*, vol. 3, no. 2, pp. 105–120, 2017.

[21] L. Demidova, E. Nikulchev, and Y. Sokolova, "Big data classification using the SVM classifiers with the modified

Research Article

GeneMiner: A Classification Approach for Detection of XSS Attacks on Web Services

Charu Gupta ⓘ**, Rakesh Kumar Singh, and Amar Kumar Mohapatra**

Department of IT, Indira Gandhi Delhi Technical University for Women, Delhi, India

Correspondence should be addressed to Charu Gupta; charugupta@igdtuw.ac.in

According to OWASP 2021, cross-site scripting (XSS) attacks are increasing through specially crafted XML documents. The attacker injects a malicious payload with a new pattern and combination of scripts, functions, and tags that deceits the existing security mechanisms in web services. This paper proposes an approach, GeneMiner, encompassing GeneMiner-E to extract new features and GeneMiner-C for classification of input payloads as malicious and nonmalicious. The proposed approach evolves itself to the changing patterns of attack payloads and identifies adversarial XSS attacks. The experiments have been conducted by collecting data from open source and generating various combinations of scripts, functions, and tags using an incremental genetic algorithm. The experimental results show that the proposed approach effectively detects newly crafted malicious XSS payloads with an accuracy of 98.5%, which is better than the existing classification techniques. The approach learns variations in the existing attack sample space and identifies the new attack payloads with reduced efforts.

1. Introduction

Web services provide a solution for information exchange over web applications developed on varied platforms [1, 2] and facilitate integration and interoperability among heterogeneous software applications using XML (eXtensible Markup Language) [3, 4]. Many available protocols and standards provide security to web services, but attacks on web services are continuously increasing [1]. According to OWASP [5], 94% of web applications are tested for one or more forms of injection attacks. The number of cross-site script (XSS) injection attacks has increased from 470 in 2011 to 22,000 in April 2022 [6], as shown in Figure 1. The injection attacks occur due to the attacker's insertion of malicious characters and strings in XML documents [1, 7]. The attacker crafts the payload in such a manner to bypass the existing filters. The injection attacks have caused disclosure, distortion, disruption, and destruction of sensitive information and defacement of websites [1]. The impact of injection attacks is scored 7.25 on an average on a scale of 10 in terms of severity [8].

Adversarial attack payload refers to a newer combination of malicious characters, strings, tags, and scripts. Attackers deceive the available security standards, mitigation approaches, security testing, filters and firewalls, intrusion detection, and prevention systems by injecting a malicious adversarial payload. Suppose an algorithm is trained to detect payloads consisting of a script and alert tags. In that case, malicious attack payloads easily deceive the trained algorithm by using newer functions and obfuscation techniques. For example, techniques trained for detecting the payload <script> alert("XSS") </script> are easily deceived by newer payload onload = alert(/hacked/) by using *onload* and obfuscation of *alert*.

Machine learning (ML) [9–13] and neural network (NN) [14–17] techniques have recently been used to detect XSS vulnerability in source code. Genetic algorithms and fuzzy inference are combined with ML and NN to generate a large number of permutations and combinations of payloads using the existing datasets to mitigate the newer adversarial attacks [18–21]. Existing algorithms are limited by training on a fixed number of features and datasets. Further, the authors observed the following issues in the existing techniques and approaches:

(1) Existing models are nonadaptive and do not fit into a dynamic and ever-changing real-world environment

where attackers generate malicious payloads, each with a different combination and unique enough to escape the filters. The malicious attack payloads have a mix of encoded characters, case mixtures, recursively nested keywords, blank spaces, and tags to bypass filters.

(2) With the ever-changing diversity in injection payloads and increase in feature set, the training algorithm becomes multidimensional, increasing the computational complexity manifolds.

In this paper, the authors propose an approach, Gene-Miner, that identifies known XSS attacks and can detect the unknown and newer attack patterns. The proposed approach has utilized the properties of the incremental genetic algorithm to detect the ever-evolving heterogeneous and changing XSS payloads. The key contributions and innovations of the paper are as follows:

(1) The proposed approach evolves to detect the adversarial malicious payloads consisting of newer functions, features, combinations, encoded characters, and obfuscation techniques.

(2) The approach is optimized to search in the large multidimensional repositories by reducing it to single-dimensional search space.

(3) The approach will learn variations in the existing attack sample space and evolve to the new attack environment consisting of a newer feature set with reduced efforts.

(4) The authors conducted experiments with 1,60,264 records for performance evaluation of the newly added feature including comparison with other widely used machine learning and neural network models.

The rest of the paper is organized as follows. In Section 2, related work is discussed. Section 3 contains the preliminaries to understand the proposed approach. In Section 4, the proposed XSS attack detection model—GeneMiner—is explained. The results of the experiments and their discussions are given in Section 5. In Section 6, conclusions and future work are discussed.

2. Related Work

NVD [8], Common Weakness Enumeration [22], and Common Attack Pattern Enumeration Catalog [23] enlisted that the most common attacks are XSS attacks via crafted XML documents [8]. XML cross-site scripting attack is carried out by manipulating the logic of XML-based web services. Hyper Text Markup Language (HTML) tags, JavaScript functions, and other characters are injected into XML messages to create an XSS attack. XSS attacks are also made by injecting parts of the attack vector in two or more input fields such that it bypasses the filter [24]. These XSS attacks on web services are capable of triggering phishing attacks, cookie theft, denial of service (DoS) attacks, distributed denial of service (DDoS) attacks, XSS worms, and browser screenshot theft.
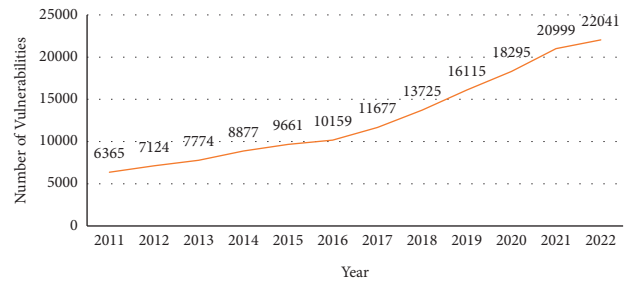
Figure 1: Increasing trend of XSS attacks.

2.1. Detection of XSS Attacks Using Neural Network.
Fang et al. [14] and Lei et al. [15] extracted six categories of features using word2vec and trained the model using LSTM recurrent neural network. The method DeepXSS [14] included a decoder to detect obfuscated malicious payloads. The network detection model proposed by [15] was based on long short-term memory (LSTM) to extract abstract features of XSS attacks. Lei et al. [15] added a hidden layer of the attention mechanism to extract more relevant information to improve the classification using the recurrent neural network.

Melicher et al. [17] trained deep neural network using taint tracking methods to predict the vulnerability of payloads by analyzing JavaScript functions. Liu et al. proposed an approach GraphXSS, for the detection of XSS attacks, which converted an XSS payload into a graph of interconnected words and characters. Fang et al. [20] combined the techniques of the deep neural network with reinforcement learning to detect the adversarial samples of XSS attacks. The detection approach generates adversarial samples and retrains the model to optimize the classification of XSS attacks.

Liu et al. [25] expressed the relationship between words and characters using the word2vec tool released by Google [26]. Liu et al. trained the model using a 2-layer and 6-layer graph convolutional network to detect XSS attacks and showed that accuracy increased with the number of layers in the detection model.

Wang et al. [27] performed experiments on 27000 samples of malicious JavaScript code and trained them using one hidden layer of stacked denoised auto-encoders forming a neural network. The accuracy of detection was 94.82%. However, the approach took longer time for training using the neural network technique. Mokbal et al. [28] designed three models for detecting XSS attacks. Raw data were collected using the first model [28] of random crawling. The second model, built on a neural network, extracted the features from raw data [28]. The third model used artificial neural network (ANN)-based multilayer perceptron (MLP) to classify and predict malicious XSS payloads [28].

2.2. Detection of XSS Attacks Using Machine Learning Algorithms.
Krishnaveni and Sathiyakumari [10] conducted experiments of 500 Uniform Resource Locators (URLs) by extracting script, applet, and DOM features to classify XSS attack payload using Naïve Bayes, decision tree, and

multilayer perceptron techniques. Vishnu and Jevitha [11] extracted 12 features from URL and JavaScript from 43,579 URL instances and classified XSS attack payload using support vector machine (SVM), Naïve Bayes, and decision tree algorithms.

Rathore et al. [29] extracted HTML and URL features from social network services and applied ten machine learning classifiers to detect an XSS attack on webpages of social network sites. The machine learning classifiers were applied on a dataset of 1000 webpages containing 400 malicious and 600 benign webpages collected from XSSed, Alexa, and Elgg sources.

Zhang et al. [30] extracted features from XSS payloads using word2vec and trained the dataset using two unsupervised clustering techniques, Gaussian mixture models (GMMs). Zhang et al. [30] built two GMMs for detecting XSS in web request and web response packets. The XSS payloads are distinguished as two clusters with two different Gaussian functions characterized by the mean and covariance of the data points in a dataset. Any addition of payloads with different feature sets will require retraining to calculate the mean and covariance of all data points again.

2.3. Detection of XSS Attacks Using Genetic Algorithm.

Khan and Motwani [31] proposed a signature-based intrusion detection of XSS attacks using a genetic algorithm with binary-coded eleven-bit chromosomes. The approach considered the count of suspicious HTML characters and words in the first three bits and the presence of encoded characters in the fifth bit, the following three bits stored the count of suspicious script characters, and the last four bits specified the class of XSS attack.

Suleman and Awan [32] applied genetic algorithm to optimize the machine learning algorithms by reducing the features selected and showed that the accuracy rate of detection of Naïve Bayes, ID3, K-nearest neighbor (KNN), decision tree, and random forest algorithms increased from 76% to 94.99% by applying genetic algorithm for feature selection. Suleman and Awan [32] selected two to eight best features out of 15 features by applying genetic algorithms to detect XSS attacks and optimized detection rate.

Tariq et al. [33] used 30 features proposed by Zhou and Wang [16] and applied a basic genetic algorithm to detect malicious XSS payloads. The model proposed trained dataset on the number of features occurring in the training dataset, and the accuracy rate was as low as 5.78% for a feature "confirm" to 69.60% for feature "script." Tariq et al. [33] added the reward policy of reinforcement learning to the detection model by updating the novel payloads to the training data to improve the accuracy rate.

2.4. Security Testing and Static Code Analysis for XSS Attacks.

Gupta and Gupta [7] proposed XSS SAFE, an automated framework for detecting and mitigating XSS attacks. The JavaScript functions were repeatedly injected into the sanitization modules of the source code, and rules were generated to mitigate the attacks. The JavaScript functions and characters were encoded using different encoding techniques and injected to generate detection rules. Gupta and Gupta [7] tested the framework on five Java Server Pages and reported a zero false-positive rate.

Hydara et al. [18] converted the source code of web applications built on Java into control flow graphs and detected the hidden vulnerabilities in control flow paths using genetic algorithms. Taint analysis was done to discover the paths of data insertion through users in five Java-based web applications with 3000 to 80000 lines of code. A genetic algorithm was then used to pass generated XSS input data to identify paths vulnerable to XSS attacks through the input points.

Ahmad and Ali [19] generated XSS test cases using a genetic algorithm and performed dynamic security testing by identifying multiple execution paths in the static code. Ahmad and Ali tested for three types of XSS vulnerability in three small web applications written in PHP. The genetic algorithm's fitness function is minimized to ensure no node is missed in a path testing of a program.

Zhang et al. [21] generated adversarial XSS attacks using the MCTS algorithm and constructed a generative adversarial network (GAN) to detect many XSS payloads.

Jan et al. [24] proposed an automated testing methodology for identifying XML injection attacks. The method randomly generates test data by using five mutation operators. The mutation operators are designed to add randomly selected XML meta-characters, remove quotes from randomly selected attributes, add a closing tag, replicate element, and replace a random element. The method showed that 78.86% of test cases generated bypassed the XML gateway filters.

Duchene et al. [34] detected only reflected XSS (type 1) attacks in web applications by generating malicious input loads using fuzzy inference and genetic algorithms. Duchene et al. [34] considered seven attributes for detecting XSS injection, namely, HTML spaces, attribute delimiter, tag limiter, JavaScript code, URL, and regular text. In their later work, Duchene et al. [35]and Kameleon Fuzz detected and tested for stored XSS attacks (type 2) by generating malicious payloads using fuzzing and genetic algorithm.

Salas and Martins [36] combined penetration testing and fault injection methods to detect XSS vulnerabilities in SOAP UI and XML messages. Salas and Martins [36] showed that the impact of XSS attacks on web services is reduced from 42% to 36% through web service security tokens.

Sahu and Tomar [37] developed a static code analysis tool for assisting web developers in writing secure code. The tool detected hidden code vulnerabilities based on three defensive programming principles. The tool ensures that all unexpected conditions are handled, external data are properly sanitized, and error messages containing technical information are disabled after development.

Liu et al. [38] introduced randomness in the initial payload and the generated high-quality vector using fuzzy inference and operators of genetic algorithm to detect XSS vulnerability in webpages. The attack payloads were generated by replacing and inserting sensitive words, characters, events, functions, blank characters, comments, tags, confusing code, and changing case structure at different location of the XSS attack payload.

2.5. Recent Surveys for XSS Attack Process, Detection, and Mitigation in Web Services. In the literature, several techniques are available that perform security testing and detection of input payloads for identifying the existence of unwanted strings and characters in the XML message packet that trigger an XSS attack. Research and review articles about various techniques for security testing and detection of XSS attacks on web services, XML messages, and web applications from the year 2010 to 2022 are studied. Gupta and Gupta [7] presented a survey of ten detection and protection techniques for XSS attacks on the server and browser sides. Gupta and Gupta [7] compared the techniques on five parameters: location of the exploit—server or browser; discovery site; technology for detection; type of XSS attack detected; and modification site—browser, server, or code. Chen et al. [13] discussed the advantages of using machine learning techniques to detect XSS attack. Chen et al. [13] discussed the limitations of whitelisting approach in detection of XSS attack. The library of filtering rule in whitelisting approach is difficult to update with the fast changing nature of XSS attack payload.

The taxonomy of XSS attacks and worms is given by Liu et al. [39]. The authors presented a state-of-the-art process to detect whether a webpage is vulnerable to XSS attack or not. Liu et al. [39] categorized defensive techniques to prevent XSS attacks into black-whitelisting, string injection, static analysis, taint flow analysis, and machine learning [39].

Mishra and Gupta [40] surveyed eleven techniques for detection and mitigation of XSS attacks including the attack via cookie theft. The authors [40] also presented an approach using cryptography techniques to prevent cookie theft through XSS payloads in web services.

Rodriguez et al. [41] analyzed more than fifty review and research articles to analyze various detection and mitigation techniques of XSS attacks. Mitigation techniques for XSS attacks include laboratory activity emulation, simulation, honeypots, IDS Snort, content analysis, rule filters, cache test, text filter, content security policy, string and URL analysis, XML, input validation, filtering patterns, attack modeling, web scanner, black box testing, defensive programming, concolic test, proxy coding of alphabets, cookies analysis, and session authentication [41].

The security of web services, SOAP UI, and XML messages can be validated through testing in the static and dynamic phases [36]. The static phase of fault detection uses code inspection, static vulnerability analysis, or theorem proof without running the system. The dynamic phase of testing searches for vulnerabilities by sending attack payloads in the request or response message. Other forms of commercial and open-source vulnerability scanners (VS) (like HP Web Inspect, BIXSAN, PathCutter, FLAX, SWAP, IBM Rational AppScan, WSDigger, WebScarab) have limited low coverage of existing vulnerabilities and a high percentage of false positives [36, 41].

Sarmah et al. [42] analyzed several XSS filters, e.g., XSS Auditor, XSS Filter, NoScript, IE8, XSS-immune, XBuster, and Rule-Based, in different web browsers. These filters use the techniques of exact or approximate string matching, string comparison, or regular expressions. However, newer and sophisticated XSS payloads and worms like the Facebook XSS worm bypass such filters. Moreover, most ML techniques for detecting malicious XSS payloads cannot detect DOM-based attacks [42].

2.6. Comparative Analysis of Recent XSS Detection Approaches. Web applications and other detection techniques undergo frequent changes with increased dependencies in the source code as the newer attack pattern is detected. These frequent changes made in the web application and security mechanisms incorporate new challenges for testing and removing the bugs [43–45]. The existing techniques in the literature require a complete restructuring of the algorithm, changing the source code [46] and retraining of the training dataset [47]. Moreover, with an increase in the size of web applications, usage of code confusion, and dynamic code generation, the various static, dynamic, and hybrid techniques of detection of XSS attacks also become more complicated [39]. The efforts required to detect changing XSS attack payloads increase exponentially. Table 1 summarizes the techniques and approaches available in the literature.

The available approaches for detecting and classifying cross-site script attacks on web services are deceived with a newer combination of malicious characters, strings, tags, and scripts commonly referred to as adversarial attacks. A fixed number of features limit existing techniques. Further, these techniques perform with lower accuracy and higher complexity as many identifying features in the dataset are added. These models also do not update the dataset with the values of newly discovered features, thereby increasing data with missing values. The problem of missing values in the existing data leads to loss of information, biased learning, and reduced statistical power.

3. Preliminaries and Terminologies

In this section, preliminaries and terminologies necessary to understand the proposed approach have been discussed.

The incremental genetic algorithm (IGA) simulates the evolving nature of attack payloads by incorporating mutation and crossover operations in the feature set [47, 48]. Newly discovered features are added in descending order of their classification accuracy to the existing feature set. The payload population with the newly added feature is then evolved using mutation and crossover. The process of mutation and crossover provides an incremental and recursive learning environment that is highly efficient and adaptable to a training set with a large dimensional space with an improved classification rate. The algorithm projects the problem of searching for a solution in a large multidimensional space by projecting it into a single-dimensional space. The algorithm reduces the cost and effort of deriving classification rules.

The IGA consists of a population with a set of chromosomes which in turn are composed of genes. The algorithm works by computing a fitness function, selection of parents, formation of the child using operations of crossover

TABLE 1: Summarized limitation of recent XSS detection and testing approaches.

Reference	Number of features	Type of XSS attack detected	Number of malicious instances in dataset	Approach	Limitations
[10]	04	R,S,D	500	ML	The dataset tested is very low in number.
[11]	12	R,S,D	43,579	ML	URL and script features were extracted. Polyglots and other complex features were not considered. Decoding of URLs was not considered
[14]	07	R,S,D	33,426	LSTM, DL	XSS attacks due to polyglots, angular JS, and CSS-style tags could not be detected. The detection algorithm lacks optimization as the number of features increases
[15]	07	R,S,D	32,168	LSTM, RNN	
[20]	07	Adversarial	33,426	LSTM, DNN, RL	
[28]	41	R, S	38,569	MLP	
[29]	25	R, S	400	ML	
[30]	10	R	35,884	GMM	DoM and stored XSS attacks were not explored
[16]	30	R,S,D	16,151	Bayesian network and EL	Attacks due to polyglots, JS, file upload, and CSS could not be detected
[17]	10	D	1,80,000	DNN	Reflected and stored XSS was not detected
[25]	07	R, S	5000	GCN	DOM type of XSS attack could not be detected
[31]	02	R,S,D	—	Signature and entropy	The detection is based on several script and URL features present in the payload. The exact pattern, sequence, and combination are not considered
[32]	15	R	11054	GA for feature selection and ML for detection	Only URL-based features were selected to detect website phishing attacks
[33]	30	R,S,D	50540	GA and RL for increasing accuracy	Polyglots and other complex XSS payloads could not be detected. Benign payloads in the dataset are considered devoid of all 30 features, while in the actual scenario, an input payload may have a few features but still not malicious. The accuracy was as low as 76% using GA, which increased after applying RL
[34, 35]	08	R, S	—	FL and GA	Attack grammar requires manual updating of the feature set. Very few attack grammar rules are generated, limiting fault detection
[27]	15	R, S	14783	Stacked denoised auto-encoder DNN	Only script tags and functions triggering XSS attacks were considered. Long training time is another limitation
[42]	21	R, S	289	ML	DOM-based XSS attack could not be detected. The training set consists of a low number of malicious instances

R: reflected XSS; S: stored XSS; D: document object model XSS; DL: deep learning; ML: machine learning; FL: fuzzy logic; RL: reinforcement learning; GA: genetic algorithm; RNN: recurrent neural network; MLP: multilayer perceptron; DNN: deep neural network; EL: ensemble learning.

and mutation, eliminating the least fit individual, and including the more fit individual. The process is repeated in a continuously evolving population until a condition is met that terminates the process.

Gene: A gene is a feature that characterizes an attack payload. A set of genes is combined into a string to form a chromosome. A vector $g = \{g_1 g_2 g_3, \ldots, g_n\}$ is a one-dimensional row vector consisting of all features identifying an injection attack payload.

Chromosomes: A chromosome, often called a "genotype," is a collection of features or attributes that provide a solution to a problem being solved through a genetic algorithm. A vector $c = \{c_1 c_2 c_3, \ldots, c_n\}$ is a one-dimensional row vector that consists of feature value identifying an injection attack payload. The value of a feature may be continuous real values or binary values.

For an attack payload, A, $c = \{c_1 c_2 c_3, \ldots, c_n\}$ is a chromosome such that

$$\forall g_i \in g,$$
$$c_i = x \text{ where } x, \text{ where } R \text{ is a set of real values,} \quad (1)$$
$$c_i = 0, \quad \text{otherwise.}$$

Population: A population is an array of chromosomes or individuals. As the new features are added, mutation and crossover are performed, and the initial population gets updated and converges to a solution recursively.

$$\text{Population } P = \begin{matrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{i1} & c_{i2} & \ldots & c_{in} \end{matrix}. \quad (2)$$

Mutation: The small deviation in the feature value is termed as "mutation" in the chromosome. The mutation is

done by changing the value of one or more features within a chromosome. For an attack payload, A,

$$\text{original chromosome } cx = \{c_1 c_2 c_3, \ldots, c_n\},$$

$$\text{mutated chromosome } cy = \{c_1' c_2' c_3', \ldots, c_n'\}, \quad (3)$$

where c' represents changed value for a gene g_i,

Crossover: The crossover is an operation where features of two-parent chromosomes are swapped to form new child chromosomes. The swapping of features may be done at a single point and multiple points, or in a random manner. The process of crossover generates a new set of child chromosomes.

$$\text{Parent1 chromosome } cx = \{c_1 c_2 c_3 c_4, \ldots, c_n\},$$

$$\text{Parent2 chromosome } cy = \{c_1' c_2' c_3' c_4', \ldots, c_n'\},$$

$$\text{After crossover at position 2,} \quad (4)$$

$$\text{Child 1 chromosome } cx' = \{c_1 c_2 c_3' c_4', \ldots, c_n'\},$$

$$\text{Child 2 chromosome } cy' = \{c_1' c_2' c_3 c_4, \ldots, c_n\}.$$

Adapting to new features: The IGA used for classifying the injection attack vector consists of the formation of the initial population, integration of old and new features, and incremental evolution of new features. The current features are retained in the population, and the newer features are added to the current chromosomes. For more than one feature discovered, the accuracy rate of each of the newly discovered features is calculated and arranged in descending order. The newly discovered features are then appended to the old chromosome ruleset to update the population. A gene set $g = \{g_1 g_2 g_3 g_4, \ldots, g_n\}$, with two newer features g_{n+1}, g_{n+2}, is integrated into the gene set such that $g' = \{g_1 g_2 g_3 g_4, \ldots, g_{in} g_{n+1}, g_{n+2}\}$ and integrated chromosome $c' = \{c_1 c_2 c_3 c_4, \ldots, c_{in} c_{n+1} c_{n+2}\}$ where c_{n+1} and c_{n+2} represent the values for the new feature $g_{n+1,}$ and g_{n+2}.

Fitness Function: The fitness of a chromosome is calculated by the percentage of payloads correctly classified by the ruleset of the chromosome [47]. The global fitness of a chromosome is defined in the whole search space. While adding the new features to form a more fit chromosome, the local fitness function considers the newer payload patterns in a local region of search space. The chromosome that correctly classifies highest number of payloads in a dataset is selected for defining the rules. Such chromosomes have highest fitness function value.

$$\text{Global Fitness function f} - \text{global: } \frac{\text{Number of Correctly classified payloads}}{\text{Total Number of payloads in the dataset}}$$

$$\quad (5)$$

$$\text{Local Fitness function f} - \text{local: } \frac{\text{Number of Correctly classified newer payloads}}{\text{Total Number of newer payloads}}.$$

Termination Criteria: The evolution process is stopped if any one of the following four criteria is reached:

(i) Maximum generation limit of the population is attained.

(ii) The best fitness of the chromosome is obtained.

(iii) There is no improvement in the accuracy and fitness of the best chromosome after a specified number of generations.

(iv) Performance of validating the data is less than 10% for the last 20 generations.

Classification Mechanism: The IGA is a rule-based classification technique that assigns each instance of a payload to a class by identifying relationships among features in it. The rules are represented in the form of IF (Condition 1) & (Condition 2) &.... Condition n-THEN (Class A_i). Each rule has one or more conditions as the antecedent, and a class is identified as the consequent. The conditions are represented as if-then rules, and the values of the set of features are represented in binary form. A ruleset contains several rules, providing a solution for a classification problem. A rule defining a class is represented as follows:

$$R_i: If\left(V_{1\_\min} \leq x_1 \leq V_{1_{\max}}\right) \land If\left(V_{2\_\min} \leq x_2 \leq V_{2\_\max}\right) \ldots \land If\left(V_{n\_\min} \leq x_n \leq V_{n_{\max}}\right)$$

$$\text{THEN } y \in C, \quad (6)$$

where R_i is a rule label, n is the number of features, $x_1 x_2, \ldots, x_n$ is the feature set, C is a class identified for payload y, and $V_{i\_\min}$ and $V_{i\_\max}$ are the minimum and maximum bounds of the feature x_i, respectively [47, 48].

4. Research Methodology

This section discusses the proposed GeneMiner approach to detect malicious payloads triggering XSS attacks, as shown in Figure 2. The authors have collected the XSS payloads
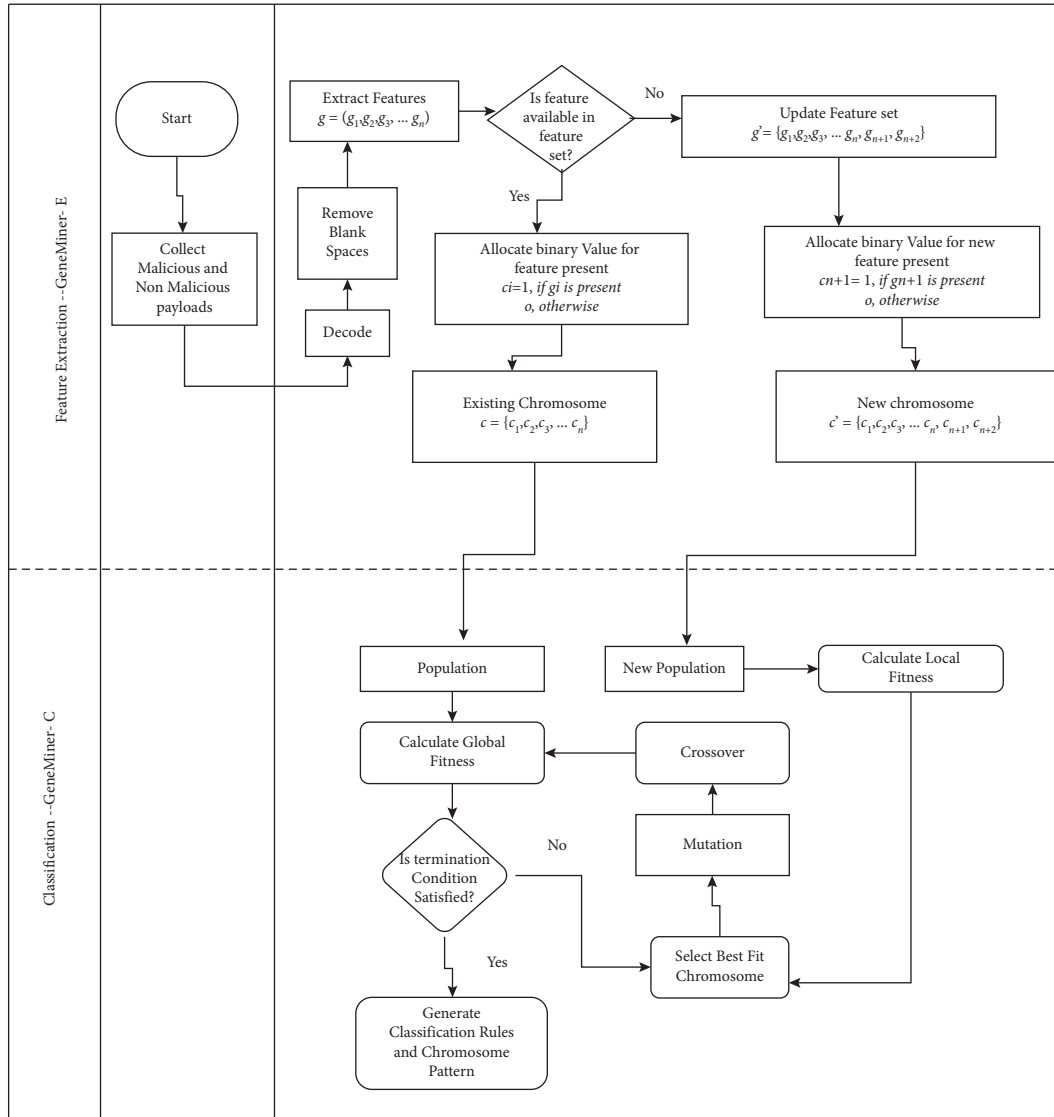
FIGURE 2: Proposed approach GeneMiner for detection of XSS attacks.

from various open-source repositories to conduct the experiments. The payload data are preprocessed, features are extracted, and the existence of a feature in the attack vector is populated in a database using the GeneMiner-E extraction model. The experiments are conducted to classify the payload as malicious or nonmalicious using the GeneMiner-C classification model.

4.1. Dataset Collection.

The authors collected the payloads from various sources for conducting the experiments to detect XSS vulnerability. The dataset contains the 6606 malicious records collected from https://github.com/payloadbox/xss-payload-list/, and 18,151 malicious and 1,35,507 nonmalicious records from https://github.com/duoergun0729/1book/tree/master/data. In total, training dataset contains 160, 264 URLs with payloads that may trigger an XSS attack. The testing dataset in the experiment is gathered from GitHub sources and security forums, which

consists of 3497 malicious and 6503 nonmalicious URLs provided by Zhou and Wang [16].

4.2. GeneMiner-E: Data Preprocessing, Decoding, and Feature Extraction.

For preprocessing of the data and extraction of features from the data, an approach, GeneMiner-E, has been proposed. The GeneMiner-E anonymizes the data by removing the domain part of the URLs and websites. Then, it decodes the encoded characters available in the payloads and cleans the data by removing blank spaces. For normalizing the data, the approach extracts the features from the payloads and prepares the final dataset by marking the feature set available in the attack vector as one and zero for the nonavailable feature.

The XSS payloads have common features of the same sensitive words and characters but differ in triggering XSS attacks. The payloads also contain characters obfuscated using encoding mechanisms such as URL encoding, ASCII

encoding, Unicode escaping, and hexadecimal encoding. These characters are decoded to original characters. For example, an attack payload may have a left-angle bracket < in hexadecimal code as %3C. The attack payload with characters 'al
er &emsp bypassed through filters as the keyword "alert" has been modified by inserting a break tag (
 and variants of blank space, namely, single space ( ), double-space ( ), four space ( ).

A payload from client to server or server to the client is classified into two classes: vulnerable and nonvulnerable. An attacker may modify the payload by injecting malicious characters, words, scripts, tags, or codes. All such malformed payloads capable of triggering an XSS attack on a webpage are labeled as malicious payloads, while the payloads that do not trigger an XSS attack are nonmalicious.

XSS payloads are identified by sensitive charact'al
er &emsp, keywords, HTML tags, scripts, redirection links, and unusual lengths of data [16, 33]. A few sensitive words and characters are alert, prompt, script, angle brackets, parenthesis, and onmouseover. The decoder module of GeneMiner also identifies the obfuscated URL modified through various encoding mechanisms, Unicode escape, inserting blank spaces, inserting invalid characters, case mixture, and other forms of malicious attack payloads [21]. The existence of these words and characters in a payload is considered a feature as it defines the characteristic nature of malicious payloads. For conducting experiments, 40 features are identified in the dataset of 1,64,204 records. Table 2 shows the list of 40 features extracted from the collected XSS payloads using GeneMiner.

4.3. GeneMiner-C: Classification Model. The training set of malicious and nonmalicious XSS payload data forms the initial population in the proposed approach. The dataset contains unique features and independent variables, while the class identified is the dependent variable. The authors constructed two datasets: one training and another testing dataset. The classification model is constructed using an incremental genetic algorithm on the training dataset and validated using the testing dataset.

The population is a binary matrix of order $m \times n$, where m is the number of URLs and n is the number of features called genes extracted from the dataset. Every row of the population matrix consists of chromosomes representing the payload as defined in

$$
\begin{aligned}
&c = \{c_1 c_2 c_3 \ldots . c_n\} \text{ is a chromosome such that} \\
&\forall g_i \in \text{gene}, g, \\
&c_i = 1 \text{ if } g_i \text{ exists in } P \text{ and,} \\
&c_i = 0, \text{otherwise.}
\end{aligned}
\tag{7}
$$

The combinations of attack payloads are generated by using a single-point crossover. The vertical dotted line in Figure 3 represents a single-point crossover. The process of mutation and crossover generates a unique

and large number of attack payloads that provides a solution to classify the payload as vulnerable or nonvulnerable.

On identifying a newer attack payload, the newly discovered features are appended to the chromosomes of the population matrix. For example, consider a payload city = test%3Cscript%3Eale%20rt%28/42873/%29%3 C/script %3E. The features extracted from the payload using the proposed GeneMiner-E are "script, left-angle-bracket, alert, open-parenthesis, close-parenthesis, right-angle-bracket." Initially, there are six features, and the chromosome corresponding to the payload will have these features marked as "1." With the addition of the new attack payload with the new feature "prompt," the number of features is increased from 06 to 07. Similarly, with the discovery of two new features, "onload" and "onmouseover," in the attack payload, the chromosome is modified by integrating the newly discovered features, as shown in Table 3. The absence of the new feature in the old payload is given a binary value of 0, and new features are appended to the modified chromosome with a value of 1.

The process of mutation and crossover continues till the termination condition is reached. A threshold of the number of generations and rules is set as a termination condition. A classification rule is obtained for each generation that classifies a certain number of chromosomes. For each classification rule, the local fitness and global fitness are calculated. The rule with the highest fitness function value, correctly classified instances, is further used in the training model to classify attack payloads. For evaluating the proposed GeneMiner, the threshold for termination condition of the proposed classification scheme has been kept at 200 generations and 100 rules by mutating one gene and single-point crossover for maximum classification rate [48].

5. Experiments and Results

The experimental environment used in this paper was run under Windows 11 operating system, equipped with a 2.6 GHz 6-core Intel Core-i7 processor and 16 GB RAM. In the experiments, the effect of adding new features and a new adversarial attack vector on the classification efficiency is analyzed on the proposed GeneMiner approach. Initially, the training dataset consists of twenty features with 44000 malicious and nonmalicious records. Newly discovered payloads are added to the initial training set in a newly discovered five feature batches. The classification accuracy of the proposed approach is compared with the popular classification techniques, i.e., Naïve Bayes (NB), random forest (RF), logistic regression (LR), support vector machine (SVM), AdaBoost, and multilayer perceptron (MLP). The results obtained are presented in Table 4.

The performance of the proposed GeneMiner is further evaluated using the accuracy, sensitivity, specificity, precision, and *F*-score metrics as defined in equations (8)–(12). The results obtained for the feature set are reported in Table 5.

TABLE 2: Features extracted for detection of XSS payload using proposed GeneMiner.

Length	Onfocus	Semicolon	Window_x	Click
Closeparen	Evalfunc	Src	Onerror	Onkeydown
Openparen	This	If_func	Sq	Onkeyup
Alert	Doc_cookie	Img	Href	Onkeypress
Close angle	Replacefunc	Onmouseover	URL	String_dot
Open angle	Try	HTTP	Prompt	Fromcharcode
Script	Catchfunc	endif_func	Iframe	Reload
Dq	expressionfunc	Formation	Change	Filters

Parent chromosome 1	11111100000000000000	000000000000011111
Parent chromosome 2	00000000000000000011	100000000000000000
Child chromosome 1	11111100000000000000	100000000000000000
Child chromosome 2	00000000000000000011	000000000000011111

FIGURE 3: Single-point crossover.

TABLE 3: Incremental updating of feature set and chromosomes using GeneMiner-E.

S.no	Attack payload	Decoded	# features	Features extracted	Initial chromosome (c_n)	Modified chromosome $(c' = c_{n+1})$	Modified chromosome $c'' = (c_{n+3})$
1	city = test%3Cscript %3Eale%20rt%28/ 42873/%29%3C/ script%3E	city = test\<script\>alert(42873) \</script\>	06	script, <, alert, ( ) >	111 111	111 111 **0**	111 111 **000**
2	city = test%3Cscript %3Eprom%20pt% 28/42873/%29%3C/ script%3E.	city = test\<script\>prompt(42873) \</script\>	07	script, <, ( ),>,prompt	—	101 111 **1** (prompt)	101 111 **100**Absence of newly discovered features is marked as 0
3	\<sCRipT < svg/ onload = '+"/ +/onmouseover = 1/ + /[∗/[]/ + alert(1)// '>	\<script < svg/ onload = onmouseover (alert1)>	09	script, <, ( ) >,onload, onmouseover	—	—	111 111 011 (onload, onmouseover)

TABLE 4: Impact of the addition of new features on accuracy of GeneMiner and other approaches.

Number of features	Naïve Bayes	RF	LR	SVM	AdaBoost	MLP	Proposed GeneMiner
20	59.25	68.41	68.48	68.48	66.48	68.48	80.73
25	66.8	64.1	64.3	69.54	68.61	69.47	91.38
30	81.6	73.7	76.1	73.19	68.61	77.30	91.45
35	88.78	69.61	98.1	87.39	87.61	87.21	96.26
40	89.05	69.47	93.76	88.98	93.49	96.80	98.50

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100, \quad (8)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100, \quad (9)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \times 100, \quad (10)$$

$$\text{Precision} = \frac{TP}{TP + FP} \times 100, \quad (11)$$

$$F - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \times 100, \quad (12)$$

where TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively.

The set of experiments conducted shows that the addition of a new feature improves the accuracy of detection of XSS attack payload using the proposed GeneMiner approach. The proposed approach has an accuracy of 98.07%, while other classifiers have accuracy as low as 69.49 for the

TABLE 5: Comparison of classification accuracy of GeneMiner with other approaches on the addition of newer attack payloads.

Approach	Accuracy	Sensitivity	Specificity	Precision	F-Score
Naïve Bayes	89.05	96.44	85.3	76.88	82.52
Random forest	69.47	44.57	82.1	55.8	49.56
Logistic regression	93.76	98.02	91.6	85.5	89.46
SVM	88.98	89.01	79.01	90.60	89.2
AdaBoost	93.49	93.50	91.07	94.10	93.30
MLP	96.80	95.29	95.29	96.80	96.30
Proposed GeneMiner	98.50	96.25	98.99	97.99	98.03

TABLE 6: Empirical comparison of GeneMiner with existing approaches.

GeneMiner proposed	GAN [21]	EL [16]	RL [20]	RF [29]	DL [27]	CNN [49]	NB [16]	SVM [16]	LR [16]	RF [16]
98.5	94.59	97.78	91.7	97.2	94.82	90.2	84.44	85.76	85.82	85.76

GAN: generative adversarial network; EL: ensemble learning; RL: reinforced learning; DL: deep learning; CNN: convolution deep neural network; NB: Naïve Bayes; SVM: support vector machine; LR: logistic regression; RF: random forest.

random forest, 89.05 for Naïve Bayes, and 93.7 for logistic regression. Logistic regression and Naïve Bayes approaches are unbalanced in classifying one class. These approaches classify negative classes more accurately with higher sensitivity, but positive classes are identified with lower accuracy. The accuracy, specificity, sensitivity, and F-Score of the proposed GeneMiner approach are high, signifying the accurate and balanced classification of vulnerable and nonvulnerable classes.

6. Empirical Evaluation with Existing Approaches

The proposed approach for the classification of payloads triggering an XSS attack is compared with the existing approaches available in the literature. The authors of [18, 19, 21, 34, 35] performed security testing using fuzzy evolutionary inference to identify XSS attacks in web applications. Khan and Motwani [31] and Suleman and Awan [32] applied a genetic algorithm for feature selection during the XSS detection methodology. Fang et al. [14, 20] performed decoding of XSS payloads and applied neural network techniques to identify malicious XSS payloads. Lei et al. [15] used the LSTM model for XSS detection, and Melicher et al. [17] applied deep neural network for detecting DOM-based XSS attacks. Liu et al. [25] applied graph convolution networks, and Abimov and Bianchi [49] used convolutional deep neural network (CNN) for the detection of XSS attacks. Zhou et al. [21, 30] applied machine learning techniques, [33] applied a genetic algorithm along with reinforcement learning to detect XSS attacks, and [50] detected reflected XSS attacks using reinforced learning. It is observed that only 06 out of 16 articles focused on evolutionary algorithms for the detection of XSS attacks, while only 01 focused on XML injection [51, 52].

The existing approaches of [14, 21, 33, 42] detect XSS payload by extracting script tags, special characters, eval function, features from URLs, and cookies but do not consider CSS style, dangling markups, and polygons. It is observed by [53] that the escape rate of XSS payloads from various detection models of machine learning and deep learning has reached 85%, and such techniques are less efficient for the detection of adversarial attacks. The proposed GeneMiner approach has extended the feature set by incorporating CSS style, dangling markups, encoded characters, and polyglots to detect malicious XSS payloads in XML messages. Table 6 shows the classification accuracy using the proposed approach and the existing approaches.

The existing approaches of XSS attack detection using generative adversarial network [30] have an accuracy rate of 94.59%, and enforcement learning [21] reaches an accuracy up to 97.78, using deep neural network [20] (with an accuracy of 91.7%). Other machine learning techniques detect XSS payloads with less than 90% accuracy. The proposed GeneMiner approach has reached the accuracy of 98.5 with 40 features detecting varied, changing, and evolving malicious XSS payloads. The existing approaches available in the literature do not evolve to changing nature and addition of newer features of XSS in attack payloads. Due to computational complexity, these approaches are also not suitable with higher dimensional feature sets. The proposed GeneMiner approach evolves itself to the modifying nature of XSS attack payloads and updates the training set incrementally with the addition of newer features. The proposed GeneMiner approach also optimizes the search space by reducing it to single-dimensional search space.

7. Conclusion and Future Work

The ever-increasing and varying nature of XSS attack payloads bypasses the filtering mechanisms of existing techniques. The XSS attack results in defacing websites, destruction of data, disruption of services, and disclosure of confidential data in web applications. This paper proposes a GeneMiner approach to detect XSS attacks in XML documents. The proposed approach evolves itself to detect malicious payloads consisting of newer functions, features, combinations, encoded characters, and obfuscation techniques. The experiments were conducted by collecting open-

source payloads and training them to detect XSS attack payloads. The proposed approach works in two steps. In the first step, the features that define the XSS attack are extracted from many payloads using the GeneMiner-E model. The newer features identified in newer malicious payloads are appended to the existing feature set. In the second step, the GeneMiner-C creates a single-dimensional ruleset to classify attack payloads. The experimental results showed that the proposed approach identifies the new adversarial attack payloads and gives a higher accuracy of 98.5%, which is better than the existing classification approaches. In the future, the authors intend to extend the proposed approach for detecting other forms of injection attacks on web services.

Data Availability

The dataset contains the 6606 malicious records collected from (1) https://github.com/payloadbox/xss-payload-list/, and 18,151 malicious and 1,35,507 nonmalicious records from (2) https://github.com/duoergun0729/1book/tree/master/data. The testing dataset in the experiment is gathered from GitHub sources and security forums, which consists of 3497 malicious and 6503 nonmalicious URLs provided by the authors of [16] Zhou, Y., amp; Wang, P. (2019).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. Gupta, R. K. Singh, and A. K. Mohapatra, "A survey and classification of XML based attacks on web applications," *Information Security Journal: A Global Perspective*, vol. 29, no. 4, pp. 183–198, 2020.

[2] R. M. Rauf, T. E. Rivaldo, and A. Amiruddin, "A review of attacks, objects, and mitigations on web services IOP Conference Series: materials Science and Engineering," *IOP Publishing*, vol. 852, no. No. 1, Article ID 012139, 2020, July.

[3] M. Kakavand, A. Mustapha, Z. Tan, S. F. Yazdani, and L. Arulsamy, "O-ADPI: online adaptive deep-packet inspector using Mahalanobis distance map for web service attacks classification," *IEEE Access*, vol. 7, Article ID 167141, 2019.

[4] N. Palsetia, G. Deepa, F. Ahmed Khan, P. S. Thilagam, and A. Pais, "Securing native XML database-driven web applications from XQuery injection vulnerabilities," *Journal of Systems and Software*, vol. 122, pp. 93–109, 2016.

[5] OWASP Foundation, "OWASP Top 10 2021," 2021, https://owasp.org/Top10/.

[6] C. Vulnerability Enumeration, "CVE," 2022, https://cve.mitre.org/.

[7] S. Gupta and B. B. Gupta, "XSS-SAFE: a server-side approach to detect and mitigate cross-site scripting (XSS) attacks in JavaScript code," *Arabian Journal for Science and Engineering*, vol. 41, no. 3, pp. 897–920, 2016.

[8] N. Vulnerability Database, "NVD," 2022, https://nvd.nist.gov/.

[9] H. B. Kazemian and S. Ahmed, "Comparisons of machine learning techniques for detecting malicious webpages," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1166–1177, 2015.

[10] S. Krishnaveni and K. Sathiyakumari, "Multiclass classification of XSS web page attack using machine learning techniques," *International Journal of Computer Application*, vol. 74, no. 12, pp. 36–40, 2013.

[11] B. A. Vishnu and K. P. Jevitha, "Prediction of cross-site scripting attack using machine learning algorithms," in *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing*, pp. 1–5, 2014, October.

[12] Y. T. Wu, S. J. Lin, H. C. Huang et al., *Detecting Cross-Site Scripting Attacks by Token Correlation*, 2010, May.

[13] X. Chen, M. Li, Y. Jiang, and Y. Sun, "A comparison of machine learning algorithms for detecting XSS attacks," *Lecture Notes in Computer Science*, Springer, in *Proceedings of the International Conference on Artificial Intelligence and Security*, pp. 214–224, Cham, 2019, July.

[14] Y. Fang, Y. Li, L. Liu, and C. Huang, "DeepXSS: cross-site scripting detection based on deep learning," in *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence*, pp. 47–51, 2018, March.

[15] L. Lei, M. Chen, C. He, and D. Li, "XSS detection technology based on LSTM-attention," in *Proceedings of the 2020 5th International Conference on Control, Robotics and Cybernetics (CRC)*, pp. 175–180, IEEE, Wuhan, China, 2020, October.

[16] Y. Zhou and P. Wang, "An ensemble learning approach for XSS attack detection with domain knowledge and threat intelligence," *Computers & Security*, vol. 82, pp. 261–269, 2019.

[17] W. Melicher, C. Fung, L. Bauer, and L. Jia, "Towards a lightweight, hybrid approach for detecting DOM XSS vulnerabilities with machine learning," in *Proceedings of the Web Conference 2021*, pp. 2684–2695, 2021, April.

[18] I. Hydara, A. B. Md Sultan, H. Zulzalil, and N. Admodisastro, "Cross-site scripting detection based on an enhanced genetic algorithm," *Indian Journal of Science and Technology*, vol. 8, no. 30, pp. 1–7, 2015.

[19] M. A. Ahmed and F. Ali, "Multiple-path testing for cross site scripting using genetic algorithms," *Journal of Systems Architecture*, vol. 64, pp. 50–62, 2016.

[20] Y. Fang, C. Huang, Y. Xu, and Y. Li, "RLXSS: optimizing XSS detection model to defend against adversarial attacks based on reinforcement learning," *Future Internet*, vol. 11, no. 8, p. 177, 2019.

[21] X. Zhang, Y. Zhou, S. Pei, J. Zhuge, and J. Chen, "Adversarial examples detection for XSS attacks based on generative adversarial networks," *IEEE Access*, vol. 8, Article ID 10989, 2020.

[22] CWE, "Common Weakness Enumeration," 2022, https://cwe.mitre.org/.

[23] CAPEC, "Common Attack Pattern Enumeration Catalogue," 2022, http://capec.mitre.org/.

[24] S. Jan, C. D. Nguyen, and L. C. Briand, "Automated and effective testing of web services for XML injection attacks," in *Proceedings of the 25th International Symposium on Software Testing and Analysis*, pp. 12–23, 2016, July.

[25] Z. Liu, Y. Fang, C. Huang, and J. Han, "GraphXSS: an efficient XSS payload detection approach based on graph convolutional network," *Computers & Security*, vol. 114, Article ID 102597, 2022.

[26] "Word2vec Google," 2022, https://code.google.com/archive/p/word2vec/.

[27] Y. Wang, W.-d. Cai, and P.-c. Wei, "A deep learning approach for detecting malicious JavaScript code," *Security and Communication Networks*, vol. 9, no. 11, pp. 1520–1534, 2016.

[28] F. M. M. Mokbal, W. Dan, A. Imran, L. Jiuchuan, F. Akhtar, and W. Xiaoxi, "MLPXSS: an integrated XSS-based attack detection scheme in web applications using multilayer perceptron technique," *IEEE Access*, vol. 7, Article ID 100567, 2019.

[29] S. Rathore, P. K. Sharma, and J. H. Park, "XSSClassifier: an efficient XSS attack detection approach based on a machine learning classifier on SNSs," *Journal of Information Processing Systems*, vol. 13, no. 4, pp. 1014–1028, 2017.

[30] J. Zhang, Y. T. Jou, and X. Li, "Cross-site scripting (XSS) detection integrates evidence in multiple stages," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019, January.

[31] S. Khan and D. Motwani, "Implementation of IDS for web application attack using an evolutionary algorithm," in *Proceedings of the 2017 International Conference on Intelligent Computing and Control (I2C2)*, pp. 1–5, IEEE, 2017, June.

[32] M. T. Suleman and S. M. Awan, "Optimization of URL-based phishing websites detection through genetic algorithms," *Automatic Control and Computer Sciences*, vol. 53, no. 4, pp. 333–341, 2019.

[33] I. Tariq, M. A. Sindhu, R. A. Abbasi, A. S. Khattak, O. Maqbool, and G. F. Siddiqui, "Resolving cross-site scripting attacks through genetic algorithm and reinforcement learning," *Expert Systems with Applications*, vol. 168, Article ID 114386, 2021.

[34] F. Duchene, R. Groz, S. Rawat, and J. L. Richier, "XSS vulnerability detection using model inference assisted evolutionary fuzzing," in *Proceedings of the 2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, pp. 815–817, IEEE, Montreal, QC, Canada, 2012, April.

[35] F. Duchene, S. Rawat, J. L. Richier, and R. Groz, "KameleonFuzz: evolutionary fuzzing for black-box XSS detection," in *Proceedings of the 4th ACM Conference on Data and application security and privacy*, pp. 37–48, 2014, March.

[36] M. I. P. Salas and E. Martins, "Security testing methodology for vulnerabilities detection of XSS in web services and WS-security," *Electronic Notes in Theoretical Computer Science*, vol. 302, pp. 133–154, 2014.

[37] D. R. Sahu and D. S. Tomar, "Analysis of web application code vulnerabilities using secure coding standards," *Arabian Journal for Science and Engineering*, vol. 42, no. 2, pp. 885–895, 2017.

[38] Z. Liu, Y. Fang, C. Huang, and Y. Xu, "GAXSS: effective payload generation method to detect XSS vulnerabilities based on genetic algorithm," *Security and Communication Networks*, 2022.

[39] M. Liu, B. Zhang, W. Chen, and X. Zhang, "A survey of exploitation and detection methods of XSS vulnerabilities," *IEEE Access*, vol. 7, Article ID 182004, 2019.

[40] P. Mishra and C. Gupta, "Cookies in cross-site scripting: type, utilisation, detection, protection and remediation," in *Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pp. 1056–1059, IEEE, Noida, India, 2020, June.

[41] G. E. Rodríguez, J. G. Torres, P. Flores, and D. E. Benavides, "Cross-site scripting (XSS) attacks and mitigation: a survey," *Computer Networks*, vol. 166, Article ID 106960, 2020.

[42] U. Sarmah, D. K. Bhattacharyya, and J. K. Kalita, "A survey of detection methods for XSS attacks," *Journal of Network and Computer Applications*, vol. 118, pp. 113–143, 2018.

[43] A. Agrawal and R. K. Singh, "Mining software repositories for revision age-based Co-change probability prediction," *International Journal of Open Source Software & Processes*, vol. 11, no. 2, pp. 16–32, 2020.

[44] A. Agrawal and R. K. Singh, "Predicting co-change probability in software applications using historical metadata," *IET Software*, vol. 14, no. 7, pp. 739–747, 2020.

[45] A. Agrawal and R. K. Singh, "Ripple effect identification in software applications," *International Journal of Open Source Software & Processes*, vol. 11, no. 1, pp. 41–56, 2020.

[46] S. Gupta and B. B. Gupta, "Cross-Site Scripting (XSS) attacks and defense mechanisms: classification and state-of-the-art," *International Journal of System Assurance Engineering and Management*, vol. 8, no. S1, pp. 512–530, 2017.

[47] H. Zhang, S.-U. Guan, and M. Xu, "Recursive learning of genetic algorithm featuring incremental attribute learning for higher dimensional classification problems," *International Journal of Machine Learning and Computing*, vol. 2, no. 6, pp. 802–806, 2012.

[48] S.-U. Guan and F. Zhu, "An incremental approach to genetic-algorithms-based classification," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 2, pp. 227–239, 2005.

[49] S. Abaimov and G. Bianchi, "CODDLE: code-injection detection with deep learning," *IEEE Access*, vol. 7, Article ID 128617, 2019.

[50] F. Caturano, G. Perrone, and S. P. Romano, "Discovering reflected cross-site scripting vulnerabilities using a multi-objective reinforcement learning environment," *Computers & Security*, vol. 103, Article ID 102204, 2021.

[51] S. Attwood, W. Li, and R. Kharel, "Evolutionary algorithms in web security: exploring untapped potential," in *Proceedings of the 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, pp. 1–6, IEEE, Porto, Portugal, 2020, July.

[52] V. S. Stency and N. Mohanasundaram, "A study on XSS attacks: intelligent detection methods," in *Journal of Physics: Conference Series* vol. 1767, no. 1, IOP Publishing, Article ID 012047, 2021, February.

[53] Q. Wang, H. Yang, G. Wu et al., "Black-box adversarial attacks on XSS attack detection model," *Computers & Security*, vol. 113, Article ID 102554, 2022.

Research Article

A Computational Intelligence Model for Legal Prediction and Decision Support

Xuerui Shang (ID)

School of Law, Shanghai University of Finance and Economics, Shanghai, 200433, China

Correspondence should be addressed to Xuerui Shang; shangxuerui@163.sufe.edu.cn

Legal judgment prediction (LJP) and decision support aim to enable machines to predict the verdict of legal cases after reading the description of facts, which is an application of artificial intelligence in the legal field. This paper proposes a legal judgment prediction model based on process supervision for the sequential dependence of each subtask in the legal judgment prediction task. Experimental results verify the effectiveness of the model framework and process monitoring mechanism adopted in this model. First, the convolutional neural network (CNN) algorithm was used to extract text features, and the principal component analysis (PCA) algorithm was used to reduce the dimension of data features. Next, the prediction model based on process supervision is proposed for the first time. When modeling the dependency relationship between sequential sub-data sets, process supervision is introduced to ensure the accuracy of the obtained dependency information, and genetic algorithm (GA) is introduced to optimize the parameters so as to improve the final prediction performance. Compared to our benchmark method, our algorithm achieved the best results on four different legal open data sets (CAIL2018_Small, CAIL2018_Large, CAIL2019_Small, and CAIL2019_Large). The realization of automatic prediction of legal judgment can not only assist judges, lawyers, and other professionals to make more efficient legal judgment but also provide legal aid for people who lack legal expertise.

1. Introduction

In traditional judicial cases, the results can be obtained only after the analysis and interpretation of relevant people, and the complexity and professionalism of legal documents are insurmountable barriers for ordinary people. In early studies, relevant people build expert systems in the judicial field to solve the problems encountered by people, but in the construction process of expert systems, legal people often need to provide a lot of rules and definitions, and the later maintenance of the system will consume a lot of time and energy. With the maturity of deep learning technology, a large number of excellent deep neural network models have been proposed, which lays a solid foundation for intelligent judicial research. Article prediction is an important subtask in intelligent justice. The article prediction task can assist judges to deal with cases; predict the law involved in the trial process, provide the basis for the judicial decision of charges, prison terms, fines, and so on; and greatly improve the

efficiency of judges to deal with cases. Legal judgment prediction (LJP) tends to enable machines to forecast the verdict of legal cases after reading the description of facts, which is an application of AI in the legal field. The realization of automatic prediction of legal judgment can not only assist judges, lawyers, and other professionals to make more efficient legal judgment but also provide legal aid for people who lack legal expertise. Recently, with the development of machine learning and natural language processing (NLP) technology, the research on the prediction of legal decisions has attracted more and more scholars' attention.

LJP research began decades ago [1–5]. At present, the legal judgment task is regarded as text classification task. Researchers have carried out researches on this task and proposed many landmark methods [6–13], mainly including the traditional machine learning method and deep neural network-based legal judgment prediction method. Most of the legal judgment methods [9, 10] based on traditional machine learning methods mark features manually based on

specific case types, which are difficult to be applied to other types of cases and have poor adaptability to scenarios. The legal judgment prediction method [11, 12] based on the deep neural network no longer relies on the manual carefully designed template but captures the context information through the convolutional neural network or cyclic neural network and makes feature representation according to the description of the case facts so as to achieve judgment prediction. Although the existing methods have achieved good prediction effect, the accuracy is still difficult to be greatly improved, mainly because the legal judgment prediction task still faces the following challenges:

(1) There are topological dependencies among legal decision sub-data sets. An LJP usually consists of detailed and complex clauses, like the relevant legal provisions involved, the crime committed by the offender, and the level of punishment (fine, prison term, etc.). Therefore, the prediction of each subtask in a legal judgment should be a whole prediction, and there is interdependence among each task [13]. Zhong et al. [13] pointed out that for human judges, there is a strict order between the sub-data sets of legal decisions.

(2) Sentence prediction based only on case description lacks prior knowledge. Compared with other sub-data sets, sentence prediction based on factual description is more challenging. In the real scene, when determining the sentence of a case, the judge will be influenced by many factors, including not only the age of the defendant described according to the facts of the case, the number of cases, but also the crime involved and other factors such as prior knowledge. Therefore, how to obtain the prior knowledge of sentence prediction under the supervision of the sub-data set information it depends on plays an important role in improving the final prediction performance. However, the existing sentence prediction task only makes prediction based on the fact description part of the case [14–16] or only considers the intermediate features of the sub-data sets it depends on in the training process as the prior knowledge of sentence prediction but lacks the process supervision of these prior knowledge [7, 9, 17].

To solve the above problems, this paper proposes a legal judgment prediction model (PS-LJP) based on process supervision. By considering the sequential topological dependence among various tasks, the prediction of relevant articles, crimes, and sentences in legal judgments is regarded as a sequential LJP problem. The process supervision information of each sub-data set is added in the training process so as to ensure the accuracy of the obtained information of the first task and then realize the effective prediction of the subsequent sub-data set. The model in this paper is based on the end-to-end recurrent neural network LJP framework. The supervision information of sub-data set labels is added to the LJP framework, and task-related features are expressed through the self-attention mechanism. The features of the prior task are taken as prior information and integrated with the features of subsequent tasks, and the result is used as the input of the subsequent prediction task to realize the decision prediction of sequential LJP.

2. Related Works

2.1. Prediction of Legal Decisions. Research on the prediction of legal judgments began many decades ago. Early studies were mostly limited by the limitations of public cases, and statistical-based methods achieved statistics for a small number of judgments rather than truly predicted [13, 18].

With the development of ML methods and NLP techniques, the legal judgment prediction task is regarded as a text classification task. Therefore, most legal judgment prediction tasks are specific to specific tasks, exploring how to extract more effective text features, which in turn achieve better crime prediction based on machine learning methods [9, 10, 17, 19, 20]. However, these traditional methods rely on hand-crafted shallow text features, with significant labor costs and poor domain adaptation, making it difficult to migrate to other scenarios.

Recently, DNN has been widely employed to [21] in the NLP field. Inspired by this, researchers try to integrate the deep neural network framework of legal knowledge for legal judgment prediction [7, 8, 13, 18, 22, 23]. For example, Luo et al. [12] adopted an NN method with attention mechanism to achieve the joint task of crime forecasting and related document extraction. Hu et al. [1–5] achieved the prediction of small sample and confusing crimes by defining 10 discriminative legal attributes. Ye et al. [6] use a sequence-to-sequence model to generate legal documents and realize crime prediction in civil cases. Zhong et al. [13] proposed the TOPJUDGE model, the first defining legal decision prediction as an LJP problem and proposed a topological LJP learning framework. However, without adding process supervision information, it is difficult to guarantee the accuracy of dependent features. Yang et al. [18] proposed an LJP learning framework that can encode multiple perspectives and multiple feedback on the dependencies and verification relationship between judgment sub-data sets, but this method is difficult to achieve end-to-end prediction and requires external auxiliary information.

2.2. Research Status of Text Classification. Traditional text classification methods are composed of feature engineering and classifier, in which feature engineering is used to extract feature information from text, and the classifier can obtain the probability distribution of each text category based on feature information. In the early text classification, bag of words, N-grams, and TF-IDF are usually used to extract features from the text, and then, support vector machine, naive Bayes, linear model, and K-nearest neighbor algorithms are combined to complete the text classification task. Word2vec tool is subsequently used to train the word vector. Each word in the text is represented as the word vector and

used as the feature vector of the text. Different machine learning algorithms are used to achieve classification prediction. The text representation obtained by traditional text classification methods is usually high-latitude sparse vector, which cannot understand the deep semantic information in the text. Deep learning text classification method uses end-to-end training neural network model and can automatically extract semantic feature information from text. Experimental results show that text classification based on deep learning is generally better than the machine learning method. Joulin [24] proposed a fast and efficient text classification model (FastText), which uses the word bag (Bow) method to represent the whole sentence and constructs the relationship between text features and text categories by linear model. FastText is not only able to achieve the accuracy of other deep learning models but also significantly faster in training tests. Yoon et al. [25] proposed text classification model (TextCNN) based on convolutional network, which includes embedding layer, convolution layer, pooling layer, and full connection layer. Word vector has two modes: static and nonstatic. By using different sizes of filter layer and pooling layer, convolution text feature extracting method will reduce the dimension of the feature vector to make it more representative. The connection layer will effectively extract the text by giving the influence of different parameters on features according to the weight of classification features. Xiang et al. [26] constructed a character-based convolutional neural network model, which takes characters as the basic unit of sentence formation and does not require word segmentation in the text classification task based on the Chinese corpus, so as to avoid the situation of poor training model due to inaccurate word segmentation. Pengfei et al. [27] coded and represented texts with recurrent neural networks, trained multiple tasks simultaneously with an LJP learning framework, and modeled different text tasks using three different information-sharing mechanisms. Zichao et al.'s [28] multilevel attention model (HAN, hierarchical attention networks) adopts the bidirectional LSTM network in the words and sentences on the two levels to form different attention mechanism at the same time and has the ability to assign large weights to important words and sentences when constructing document representations. Siwei et al. [29] proposed a recursive convolutional neural network text classification model, using recursive structure to capture context information, and the maximum pooling layer in the convolutional neural network to determine which words play a key role in text classification so as to capture key elements in the text. Conneau et al. [30] paid attention to the influence of layers of the convolutional neural network on the classification effect, and the experimental results showed that increasing the number of convolutional layers was conducive to the extraction of more comprehensive text features.

2.3. Review of Similar Case-Matching Research.

Similar case matching mainly compares the text similarity of two legal documents and selects the most similar one from the two candidate documents, which is a text similarity task in essence. However, due to the general description framework of legal documents and the professional words in the documents, higher requirements are put forward for similar case matching method. At present, most of the research on similar case matching is based on text similarity task, and the method is optimized according to the specific situation of the task. We then introduce the study of text similarity.

2.4. Review of Text Similarity Research.

Text similarity is an important task in the field of natural language processing and has very important practical significance for other tasks, such as information retrieval, reading comprehension, and abstract generation. Early researchers proposed string-based methods by analyzing literal surface information in text. Representative methods include longest common substring, edit distance, and Jaccard. These methods simply consider the composition of characters or words and can generally be used for simple text similarity calculation tasks. Salton et al. [31] proposed a vector space model (VSM), and employed it to choose an optimum indexing vocabulary for a collection of documents. The evaluation results indicate the usefulness of the model. However, the document information of different words was counted and expressed in vector form by term frequency and inverse document frequency (TF-IDF); the disadvantage of this method is that the constructed text vector is usually high dimensional and sparse. The calculation efficiency is low. Landauer et al. [32] proposed the latent semantic analysis method (LSA), whose basic idea is to reduce the dimension of the high-dimensional sparse matrix represented by text by using singular value decomposition technology, so that the data finally obtained do not have high-dimensional sparsity and can better represent text information. Hofmann [33] improved the PLSA model based on LSA, and the model used maximum expectation algorithm to calculate text topic in LSA. Blei et al. [34] proposed an LDA topic model, whose basic idea is to use three different levels of the Bayesian model to model text topic, get text topic distribution through layer upon layer traversal, and calculate text similarity value by using this distribution. In recent years, deep learning techniques have been widely used in text similarity tasks. Mikolov et al. [35] proposed the Word2Vec word vector model tool. Word2Vec contains two different modes, CBOW or CSkip, to train the language model through shallow neural networks. Each word is represented as word embedding with general semantic information. On the basis of the word vector, each word in the text is represented as a fixed dimension vector, and the similarity of the text is calculated by calculating the similarity between the word vectors. Mikolov subsequently proposed the Doc2vec method [36], which is similar to the word vector representation method trained by Word2vec. Compared with word vector representation trained by Word2vec for a single word, Doc2vec can train word vector representation for the whole document. Doc2vec can avoid semantic deviation of the whole document caused by some inaccurate word vectors, so it has more advantages in the semantic representation of the whole document. Mueller et al. [37] proposed Siamse LSTM based on twin structure, which consists of two subrecurrent neural networks. Each

recurrent neural network uses LSTM to encode and represent each text, and each text is represented as a vector of the same dimension. Finally, cosine similarity is used to calculate the similarity of two texts. Yin et al. [38] proposed a convolutional neural network (ABCNN) based on twin structure, which is used for sentence pair matching task. The model contains two subconvolutional neural networks, and an attention mechanism is added to the input layer and the convolutional layer of the convolutional neural network to better construct the semantic representation of text. Recently, Google researcher Devlin [39] proposed a pre-training model (Bert) based on bidirectional transformer encoder, which has the function of calculating the similarity of two texts. The training process includes pretraining and fine-tuning. Firstly, pretraining is carried out on large-scale unsupervised data so that the pretraining model has general semantic information. Secondly, the model is fine-tuned according to the specific task so that the model has the semantics of the specific task. The model has achieved breakthrough results in many natural language processing tasks.

This paper proposes a sequential LJP legal decision prediction model based on process supervision for end-to-end legal decision prediction. Different from the traditional LJP learning study on how to realize parameter sharing, the model in this paper not only models the sequential dependencies among various sub-data sets but also introduces a process supervision mechanism so as to realize the prediction of legal decisions by integrating prior information.

3. Data Processing

3.1. Feature Extraction Processing. Convolution neural network is one of the representative methods of deep learning technology, the convolutional neural network first began in the 1980s, and with the rise of artificial intelligence and machine computing power to ascend and the convolution neural network to get fast development, the different structures of convolution neural network are applied in computer vision, image processing, and other fields. A simple convolution neural network usually contains the input layer, convolution meter calculate layer, and pooling layer. The convolution meter calculation is the core of the CNN, which is used to extract local features of image or text information.

Researchers mainly mine text semantic features by constructing feature engineering and emotion dictionary, but they need to spend a lot of manpower on feature selection and design. In deep learning, the one-dimensional convolutional neural network CNN can obtain the local main semantic information of the text. Therefore, we used CNN network to extract features from the data in this paper. The structure of the CNN model includes three hidden layers in this paper as shown in Figure 1.

3.2. Dimension Reduction Processing. In the process of processing high-dimensional data in the model, important independent variables can be obtained by using variable selection, which can reduce the complexity of the model and ensure that the screened independent variables have a strong interpretation of the dependent variables. This method makes the model have the excellent performance of low-dimensional ideal model, so it has become a common method of scholars in various fields. The starting point of the variable selection method is summarized. First, the independent variable level is studied to test whether each independent variable can enter the model by establishing appropriate rules. These studies have formed traditional model selection algorithms based on AIC criterion, BIC criterion, RIC criterion, and so on. However, in the case of high dimensions, this kind of thinking is likely to fall into the dilemma of difficult calculation, that is, the frequently mentioned "dimension disaster" problem. PCA algorithm was proposed and widely used in various fields [40]. In our model, we will use the PCA algorithm to reduce the dimension of feature data.

4. A Computational Intelligence Model for Legal Prediction and Decision Support

In this paper, a sequential legal decision prediction model (PS-LJP) based on process supervision is proposed. The framework of the cyclic neural network which introduces process supervision is introduced.

4.1. Sequential Legal Decision Prediction Model (PS-LJP) Based on Process Supervision. This article proposes a sequence decision prediction framework based on process supervision, which consists of three layers: shared fact description coding layer, process supervision layer based on self-attention mechanism, and output prediction layer.

4.1.1. Shared Fact Description Coding Layer. In the prediction framework adopted in this paper, the fact description coding layer is the shared layer of the prediction. For the description of facts, this paper adopts BiLSTM to encode the description of facts. To input, the LSTM unit calculates through the input gate, forgettable gate, update gate, and output gate and carries out characteristic coding on the input.

At each time step $t \in [1, n]$, the LSTM unit takes x_t as the input, recalculates the storage unit c_t, and outputs the new hidden state h_t as shown in the following equations:

$$f_t = \sigma\big(W_f[x_t; h_{t-1}] + b_f\big). \tag{1}$$

$$i_t = \sigma\big(W_i[x_t; h_{t-1}] + b_i\big). \tag{2}$$

$$o_t = \sigma\big(W_o[x_t; h_{t-1}] + b_o\big). \tag{3}$$

$$\widehat{c}_t = \tanh\big(W_c[x_t; h_{t-1}] + b_c\big). \tag{4}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widehat{c}_t. \tag{5}$$
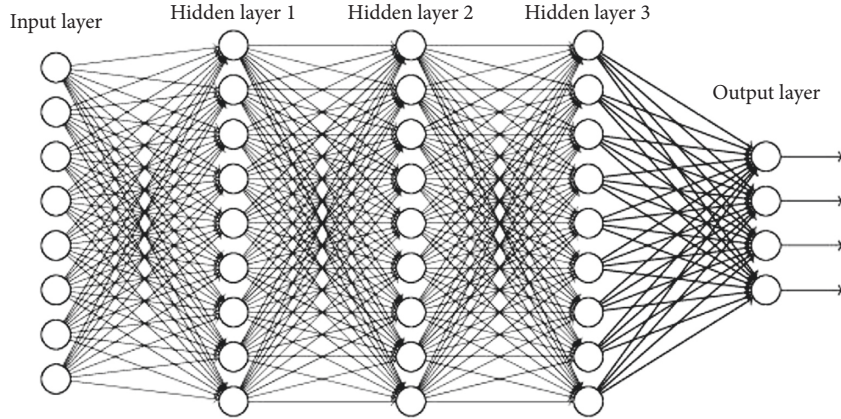
FIGURE 1: The structure of CNN model.

$$h_t = o_t \cdot \tanh(c_{t-1}). \tag{6}$$

In the BiLSTM coding process, the forward LSTM network can obtain the left-to-right feature representation of the fact description, and the backward LSTM network can encode and fuse the semantics of subsequent features. Therefore, two hidden layer feature representation sequences $\{\overrightarrow{h}_1, \overrightarrow{h}_2, \cdots \overrightarrow{h}_n\}$ and $\{\overleftarrow{h}_1, \overleftarrow{h}_2, \cdots \overleftarrow{h}_n\}$ and the two implicit states of each word are spliced together to obtain the implicit state representation $h_j = \overrightarrow{h}_j; \overleftarrow{h}_j$ and finally obtain the following equation:

$$h = BiLSTM(x, \phi). \tag{7}$$

In other words, the input vector x obtains high-level semantic output h through the coding of the BiLSTM network. Here, the dimension of the forward and backward implicit state is set as $d/2$, and the dimension of implicit state after splicing is set as d, and ϕ is the parameter involved in the above coding process.

4.1.2. Process Supervision Layer Based on Self-Attention Mechanism. The process supervision layer based on the self-attention mechanism mainly introduces process supervision to obtain effective prior task-related dependency information, thus providing an important feature guarantee for subsequent sequential LJP dependency information fusion layer extraction. In the monitoring layer of the LJP process, the BiLSTM network based on the self-attention mechanism is adopted to obtain the fact description features of each sub-data set by using sub-data set label information to supervise.

Input the output h of the shared fact description coding layer into the BiLSTM network of $|T|$ sub-data sets, respectively, and obtain the high-level semantic representation of each task:

$$h^i = BiLSTM(x, \phi_i), i = 1, 2, \cdots, |T|, \tag{8}$$

where ϕ_i is the parameter of BiLSTM corresponding to the $i - th$ sub-data set. To better capture the relevant characteristics of each task, we introduce the attention mechanism

by entering h^i into a full connection layer and mapping it between [0,1] using the softmax function.

$$a^i = softmax(W_i^p h^i + b_{ai}). \tag{9}$$

Then, the feature representation for each sub-data set classification is obtained by weighted sum based on self-attention weight.

$$\widehat{h}^i = h^i \cdot a^i. \tag{10}$$

In order to ensure that each sub-data set can obtain its related feature representation from the fact description feature in modeling, category labels of each sub-data set are added for process supervision during this part of training, namely,

$$\widehat{y}_{pi} = \text{soft max}\left(W_i^p \widehat{h}^i + b_i^p\right). \tag{11}$$

Based on the obtained prediction result $\widehat{y}_{pi}$, the minimum cross-entropy between $\widehat{y}_{pi}$ and y_{pi} is adopted.

$$L_{pi}(\widehat{y}_{pi}, y_{pi}) = -\sum_{k=1}^{|Y_i|} y_{pi,k} \log(\widehat{y}_{pi,k}). \tag{12}$$

Process supervision is introduced to ensure that it can obtain effective sub-data set-oriented feature representation, which can guarantee the accuracy of sequential LJP prediction based on the dependency feature.

4.1.3. Output Prediction Layer. Based on the feature representation h_{dj} of fusing dependent task information, it is linearly transformed and the final prediction of the task containing dependent information is realized by the softmax function so as to minimize the cross-entropy of the task containing dependent information.

$$\widehat{y}_{dj} = \text{soft max}\left(W_j^d \widehat{h}^j + b_j^d\right),$$

$$L_{dj}(\widehat{y}_{dj}, y_{dj}) = -\sum_{k=1}^{|Y_i|} y_{dj,k} \log(\widehat{y}_{dj,k}). \tag{13}$$

4.1.4. Training. For all prediction tasks, we weighted the sum of the process-supervised cross-entropy loss of each sub-data set and the final cross-entropy of the task containing dependent information so as to obtain the final loss function.

$$L = \sum_{i=1}^{|T|} \alpha_{i \cdot pi}\left(\widehat{y}_{pi}, y_{pi}\right) + \sum_{j=1}^{|T|} \beta_{j \cdot dj}\left(\widehat{y}_{dj}, y_{dj}\right), \tag{14}$$

where α_i is the weight coefficient of each sub-data set in the process of process supervision and β_j is the weight coefficient of each sub-data set in the training process containing dependent information. In practice, we keep the weight of the loss function of process supervision and the weight of the sub-data set containing dependent information consistent. We only need to obtain the weight proportion of the two parts. Here, GA is used for parameter optimization.

5. Experiment

5.1. Data Set. The law prediction task uses the China Judicial AI Challenge data set CAIL2018, which is the first large-scale Chinese data set for predicting legal decisions. The data come from real criminal cases published by the Supreme People's Court. Each sample is judged by the factual description of the case and the corresponding outcome. In actual cases, many involve more than one defendant, which will greatly increase the difficulty of legal forecasting. Therefore, this paper only retains the sample of cases composed of a single defendant. According to statistics, the distribution of law categories in CAIL2018 is quite unbalanced. Some common crimes, such as theft and intentional injury, account for a large proportion. The data of the ten laws with the highest frequency account for 79.0% of the total data, while the data of the ten laws with the lowest frequency only account for 0.12% of the total data. The imbalance of data classes in CAIL2018 is a great challenge for the prediction of low frequency and obfuscation laws. The CAIL2018 data set consists of Small_CAIL2018 and Large_CAIL2018 data sets. Small_2018 contains 196,000 instrument samples, and Large_2018 contains 1.5 million instrument samples. The division of the two data sets in the experiment is shown in Table 1. In addition, we supplemented the CAIL2019 database to further verify the validity of our proposed model.

We use accuracy (Acc.), macro average precision (MP), macro average recall rate (MR), and macro average F1 value (F1) as evaluation indexes, where the macro average indicator refers to the average of all categories of indicators.

5.2. Experimental Results and Analysis. This article evaluates the prediction effect of four sub-data sets of legal judgment prediction: CAIL2018_Small, CAIL2018_Large, CAIL2019_Small, and CAIL2019_Large. To compare the experimental results, the following three models were used as the baseline models (each baseline system was trained under the LJP frame, and the best experimental results were compared):

Table 1: Statistic information of CAIL2018 data sets.

Data sets	Small	Large
Training	150000	1450000
Test	30000	30000
Valid	16000	20000

Table 2: Prediction results of CAIL2019_Large.

Methods	Data set	CAIL2019_Large			
		Acc.	MP	MR	F1
Baselines	CNN	0.8219	0.8543	0.8836	0.8176
	HLSTM	0.8443	0.878	0.8082	0.8418
	TOPJUDGE	0.8444	0.8892	0.8015	0.8433
Ours	LJP	0.8453	0.884	0.8064	0.8436
	PS-LJP	0.8568	0.878	0.8368	0.57
	PCA-PS-LJP	0.8791	0.8786	0.8726	0.8682

(1) *CNN*. In the text, the CNN model [15] containing multiple filter lengths is used to characterize and classify the fact description

(2) *HSLTM*. Referring to the hierarchical neural network structure adopted by Yang et al. [41] in the emotion classification task, this paper adopts LSTM to encode sentence features and another LSTM to encode document features described by facts

(3) *TOPJUDGE*. Zhong et al. [13] proposed a legal decision prediction model considering topological dependency among various sub-data sets, in which the representation of fact description features was obtained by LSTM coding

Tables 2–4 list the experimental results of the pretest for the four sub-data sets. The experimental results show that the PS-LJP model adopted in this paper outperforms the baseline system in four classification indexes of four data sets, which demonstrates the effectiveness of the model presented in this paper. Tables 2–5 list the experimental results.

It can be seen from Tables 2–5 that, (1) in the four data sets, all the models perform better in the Small data set than in the Large data set, which indicates that the size of the data set will directly affect the prediction accuracy of our method; (2) in the benchmark model and our proposed model, our model outperforms the benchmark model in the four data sets. This shows that our proposed model is better than some classical benchmark models for predicting legal decisions; (3) in general, compared with the single LJP, the PS-LJP model proposed by us has better performance in four different indicators on four data sets; (4) compared with the PS-LJP model without feature data preprocessing, the PS-LJP model based on principal component analysis proposed by us achieved better results in our four data sets.

More importantly, we found that the PCA-PS-LJP model we proposed achieved the best prediction results among all the models, which will prove that the method we proposed is very effective.

TABLE 3: Results of CAIL2019_Small.

Methods	Data set	CAIL2019_Small			
		Acc.	MP	MR	F1
Baselines	CNN	0.8552	0.8765	0.8967	0.827
	HLSTM	0.8776	0.9002	0.8213	0.8512
	TOPJUDGE	0.8777	0.9114	0.8146	0.8627
Ours	LJP	0.8786	0.9162	0.8195	0.853
	PS-LJP	0.8901	0.9002	0.8499	0.8794
	PCA-PS-LJP	0.9124	0.9008	0.8857	0.8776

TABLE 4: Prediction results of CAIL2018_Large.

Methods	Data set	CAIL2018_Large			
		Acc.	MP	MR	F1
Baselines	CNN	0.8374	0.8742	0.8955	0.8361
	HLSTM	0.8698	0.8979	0.8201	0.8603
	TOPJUDGE	0.8599	0.9091	0.8134	0.8618
Ours	LJP	0.8608	0.9139	0.8283	0.8621
	PS-LJP	0.8723	0.8979	0.8487	0.8785
	PCA-PS-LJP	0.8946	0.8985	0.8845	0.8867

TABLE 5: Prediction results of CAIL2018_Small.

Methods	Data set	CAIL2018_Small			
		Acc.	MP	MR	F1
Baselines	CNN	0.8717	0.8876	0.9288	0.8488
	HLSTM	0.8951	0.9113	0.8534	0.873
	TOPJUDGE	0.8942	0.9125	0.8467	0.8745
Ours	LJP	0.8951	0.9173	0.8616	0.8748
	PS-LJP	0.9066	0.9113	0.882	0.9012
	PCA-PS-LJP	0.9289	0.9119	0.9178	0.8994

6. Conclusion

As an important subtask of intelligent justice, the study of law prediction is of great value. It can not only help legal personnel to deal with cases and improve work efficiency but also help ordinary people to understand cases and make them have certain psychological expectations for the outcome of cases. Based on the research of law prediction in intelligent justice, this paper analyzes the problems from the judicial perspective and converts them into text processing tasks. It adopts the deep learning technology to construct different neural network models and conducts training and testing on the basis of real legal data sets. In the law prediction task in this article, convolutional neural network (CNN) algorithm was used to extract text features, and the principal component analysis (PCA) algorithm was used to reduce the dimension of data features. Next, the prediction model based on process supervision is proposed for the first time. When modeling the dependency relationship between sequential sub-data sets, process supervision is introduced to ensure the accuracy of the obtained dependency information, genetic algorithm (GA) is introduced to optimize the parameters so as to improve the final prediction performance. Compared to our benchmark method, our algorithm achieved the best results on four different legal open data sets (CAIL2018_Small, CAIL2018_Large, CAIL2019_Small, and CAIL2019_Large).

Data Availability

Original data can be accessed through the following website: https://github.com/thunlp/CAIL.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] F. Kort, "Predicting Supreme court decisions mathematically: a quantitative analysis of the "right to counsel" cases," *American Political Science Review*, vol. 51, no. 1, pp. 1–12, 1957.

[2] S. S. . Nagel, "Applying correlation analysis to case prediction," *Tex.l.rev*, vol. 42, no. 7, pp. 1006–1017, 1964.

[3] R. Keown, "Mathematical models for legal prediction," *Computer Journal*, vol. 2, no. 1, pp. 829–831, 1980.

[4] J. A. Segal, "Predicting Supreme court cases probabilistically: the search and seizure cases, 1962-1981," *American Political Science Review*, vol. 78, no. 4, pp. 891–900, 1984.

[5] B. E. Lauderdale and T. S. Clark, "The Supreme court's many median justices," *American Political Science Review*, vol. 106, no. 4, pp. 847–866, 2012.

[6] H. Chen, D. Cai, and W. Dai, "Charge-based prison term prediction with deep gating network," *EMNLP*, pp. 6362–6367, 2019.

[7] H. Ye, X. Jiang, and Z. Luo, "Interpretable charge predictions for criminal cases: learning to generate court views from fact descriptions," *Computer Language*, 2018.

[8] J. Xin, Y. Hai, and Z. Luo, "Interpretable rationale augmented charge prediction system," in *Proceedings of the Coling 2018*, January 2018.

[9] C. L. Liu, C. T. Chang, and J. H. Ho, "Case instance generation and refinement for case-based criminal summary judgments in Chinese[J]," *Journal of Information Science and Engineering*, vol. 20, no. 4, pp. 783–800, 2004.

[10] C.-L. Liu and C.-D. Hsieh, "Exploring phrase-based classification of judicial documents for criminal charges in Chinese," *Lecture Notes in Computer Science*, pp. 681–690, 2006.

[11] D. M. Katz, M. J. Bommarito, and J. Blackman, "A general approach for predicting the behavior of the Supreme Court of the United States," *PLoS One*, vol. 12, no. 4, p. e0174698, 2017.

[12] B. Luo, Y. Feng, and J. Xu, "Learning to predict charges for criminal cases with legal basis," *arXiv*, 2017.

[13] H. Zhong, Z. Guo, and C. T. U, "Legal judgment prediction via topological learning," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Beijing, China, January 2018.

[14] B. Baharum, L. H. Lee, and K. Khairullah, "A review of machine learning algorithms for text-documents classification [J]," *Journal of Advances in Information Technology*, vol. 1, no. 1, pp. 4–20, 2010.

[15] Y. Kim, "Convolutional neural networks for sentence classification," *Eprint Arxiv*, 2014.

[16] D. Tang, Q. Bing, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Harbin, China, August 2015.

[17] W. C. Lin, T. T. Kuo, and T. J. Chang, "Exploiting machine learning models for Chinese legal documents labeling, case classification, and sentencing prediction," in *Proceedings of the 24th Conference on Computational Linguistics and Speech Processing, ROCLING 2012*, Taiwan, September 2012.

[18] W. Yang, W. Jia, and X. Zhou, "Legal judgment prediction via multi-perspective Bi-feedback network," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, AAAI, August 2019.

[19] N. Aletras, D. Tsarapatsanis, D. Preoţiuc-Pietro, and V. Lampos, "Predicting judicial decisions of the European court of human rights: a natural language processing perspective," *Peerj Computer Science*, vol. 2, no. 10, p. e93, 2016.

[20] O. M. Sulea, M. Zampieri, and S. Malmasi, "Exploring the use of text classification in the legal domain," in *Proceedings of the 2nd Workshop on Automated Semantic Analysis of Information in Legal Texts (ASAIL)*, London, United Kingdom, 2017.

[21] Z. K. Hu, X. Li, and C. C. Tu, "Few-shot charge prediction with discriminative legal attributes," in *Proceedings of the 27th International Conference on Computational Linguistic*, pp. 847–498, Beijing, China, August 2018.

[22] S. Long, C. Tu, and Z. Liu, "Automatic judgment prediction via legal reading comprehension," *Chinese Computational Linguistics*, vol. 11856, pp. 558–572, 2018.

[23] M. R. S. Marques, T. Bianco, and M. Roodnejad, "Machine learning for explaining and ranking the most influential matters of law," in *Proceedings of the the Seventeenth International Conference*, June 2019.

[24] A. Joulin, E. Grave, and P. Bojanowski, "Bag of tricks for efficient TextClassification," *arXiv*, pp. 427–431, 2016.

[25] K. Yoon, "Convolutional neural networks for sentence classification," in *Proceedings of the EMNLP*, pp. 1746–1751, August 2014.

[26] X. Zhang, J. Zhao, and Y. Lecun, "Character-level convolutional networks for text classification," in *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 1–9, MIT Press, 2015.

[27] P. Liu, X. Qiu, and X. Huang, "Recurrent neural network for text classification with LJP learning," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 2873–2879, AAAI Press, 2016.

[28] Z. Yang, D. Yang, and C. Dyer, "Hierarchical attention networks for document classification," in *Proceedings of the NAACL*, pp. 1480–1489, 2016.

[29] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification," *AAAI*, pp. 2267–2273, 2015.

[30] A. Conneau, H. Schwenk, and L. Barrault, "Very deep convolutional networks for text classification," 2016, http://arxiv.org/abs/1606.01781.

[31] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[32] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, vol. 41, no. 6, pp. 391–407, 1990.

[33] T. Hofmann, "Probabilistic latent semantic indexing[C]," *Research & Development in Information Retrieval*, vol. 42, no. 1, pp. 50–57, 1999.

[34] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation[J]," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[35] T. Mikolov, K. Chen, and G. Corrado, "Efficient estimation of word representations in vector space[J]," 2013, http://arxiv.org/abs/1301.3781.

[36] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proceedings of the International conference on machine learning*, pp. 1188–1196, PMLR, 2014.

[37] J. Mueller and A. Thyagarajan, "Siamese recurrent architectures for learning sentence similarity," in *Proceedings of the thirtieth AAAI conference on artificial intelligence*, pp. 2786–2792, AAAI, 2016.

[38] W. Yin, H. Schütze, B. Xiang, and B. Zhou, "ABCNN: attention-based convolutional neural network for modeling sentence pairs," *Transactions of the Association for Computational Linguistics*, vol. 4, pp. 259–272, 2016.

[39] J. Devlin, M. W. Chang, and K. Lee, "Bert: pre-training of deep bidirectional transformers for language understanding[J]," 2018, http://arxiv.org/abs/1810.04805.

[40] S. Deng, C. Wang, Z. Fu, and M. Wang, "An intelligent system for insider trading identification in Chinese security market," *Computational Economics*, vol. 57, no. 2, pp. 593–616, 2020.

[41] Z. Yang, D. Yang, and C. Dyer, "Hierarchical attention networks for document classification," in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Redmond, 2016.

Research Article

Sustainable Smart Industry: A Secure and Energy Efficient Consensus Mechanism for Artificial Intelligence Enabled Industrial Internet of Things

A. Sasikumar,[1] Logesh Ravi,[2] Ketan Kotecha [iD],[3] Jatinderkumar R. Saini,[4] Vijayakumar Varadarajan,[5,6] and V. Subramaniyaswamy [iD][7]

[1]*Department of Electronics and Communication Engineering,*
 Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India
[2]*SENSE, Vellore Institute of Technology, Chennai, Tamilnadu, India*
[3]*Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed University), Pune, India*
[4]*Symbiosis Institute of Computer Studies and Research, Symbiosis International (Deemed University), Pune, India*
[5]*Ajeenkya DY Patil University, Pune, India*
[6]*School of Computer Science and Engineering, University of New South Wales Sydney, Kensington, NSW, Australia*
[7]*School of Computing, SASTRA Deemed University, Thanjavur, India*

Correspondence should be addressed to Ketan Kotecha; head@scaai.siu.edu.in and V. Subramaniyaswamy; vsubramaniyaswamy@gmail.com

In recent years, the Internet of Things (IoT) has been industrializing in various real-world applications, including smart industry and smart grids, to make human existence more reliable. An overwhelming volume of sensing data is produced from numerous sensor devices as the Industrial IoT (IIoT) becomes more industrialized. Artificial Intelligence (AI) plays a vital part in big data analyses as a powerful analytic tool that provides flexible and reliable information insights in real-time. However, there are some difficulties in designing and developing a useful big data analysis tool using machine learning, such as a centralized approach, security, privacy, resource limitations, and a lack of sufficient training data. On the other hand, Blockchain promotes a decentralized architecture for IIoT applications. It encourages the secure data exchange and resources among the various nodes of the IoT network, removing centralized control and overcoming the industry's current challenges. Our proposed approach goal is to design and implement a consensus mechanism that incorporates Blockchain and AI to allow successful big data analysis. This work presents an improved Delegated Proof of Stake (DPoS) algorithm-based IIoT network that combines Blockchain and AI for real-time data transmission. To accelerate IIoT block generation, nodes use an improved DPoS to reach a consensus for selecting delegates and store block information in the trading node. The proposed approach is evaluated regarding energy consumption and transaction efficiency compared with the exciting consensus mechanism. The evaluation results reveal that the proposed consensus algorithm reduces energy consumption and addresses current security issues.

1. Introduction

The digitization transition gives expressive possibilities for the industry to grow creative and changing economic models and complex circular distribution networks. However, the information technology and transmission sector have a small impact on the environment; such a conversion has significant consequences for sustainability.

It is vital to supply solutions in a resilient and comprehensive manner throughout their entire life cycle to meet the milestones set forth by the industrial revolution for sustainable development [1] and achieve the circular economy's goals. Three fundamental innovative models enable the long-term digitization of a smart circular economy: industrial IoT, edge-based computing, and artificial intelligence (AI).

The introduction of the context of big data and two dominant digital innovations, such as machine learning and the Internet of Things, has recently been experienced around the world. Whereas the Internet of Things establishes a network of interlinked systems, machine learning (AI) allows machines to mimic cognitive abilities. AI and the IoT can work together to allow a new potential technology called Artificial Intelligence of Things (AIoT). In general, AIoT intends to make IoT production effective, increase human-machine interactions, and improve data gathering and analysis capabilities. These innovative techniques are developed through intelligent system advances in hardware (e.g. edge devices and accelerators) and software (e.g. RTOS, digital twin, deep learning architecture); AIoT is becoming the real-time application. In recent times, AIoT has been adopted in many application areas such as smart home automation [2], industry automation, and smart cities [3].

AIoT, as an AI-enabled framework, follows the standard training and inference approach [4] depicted in Figure 1. In the first stage, AI models utilise various machine learning techniques for training the data set. In the deployment of AIoT, training data have frequently gathered various data from IoT devices. Models are developed in the second step to arrive at conclusions from specific information. The two processes are commonly referred to as model development and Inferencing. A fundamental difficulty in AIoT is that the development of the model algorithm stage necessitates a large quantity of data and processing capabilities to produce the best AI models [5]. Yet, most IoT devices lack the necessary storage resources due to different constraints.

With the development of smart sensor technologies to integrate AI-based systems deployed in real-time applications, all data start from the era of the Industry 4.0 revolution [6]. Smart sensors are a topic that contributes to the enhancement of production and increased turnover in a variety of industries [7]. These advantages have been proven, especially when the technology available on the market is used effectively. However, AI applications can be harmful in some situations, causing major problems for the company in question. Furthermore, sensors can respond differently in different environments. They may give data of varying quality, which might misidentify the model decision and result in categorization failures if the model is not sufficiently stable. A significant effort and high costs are associated with an AI-based system developed to solve a single classification challenge, and a single misclassification scenario is costly.

The disadvantages of misclassification differ from one area to the next, based on a specific domain. On the one side, in the medical field, when a computerized diagnosis suggests that a person is sick, but in reality his health is not that poor [8]. Later, a doctor can verify and discover that the patient is, in fact, healthy. In the opposite situation, failing to recognize a sick patient and allowing him or her to continue without treatment is extremely harmful. Human safety is taken into account here; hence, high classification accuracy of greater than 99.99 per cent is essential. On the other hand, we permit marginally greater categorization errors for most industrial settings that do not endanger people's safety.
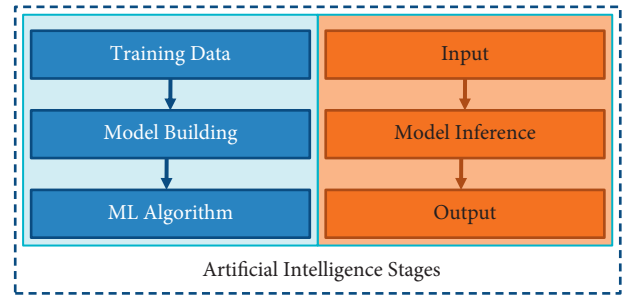
FIGURE 1: Artificial Intelligence two-stage process.

Based on various research studies in the agricultural field, the work of Xiong et al. [9], as well as Wossen et al. [10], validated that the cost of a misclassification error varies based on whether it is a false-negative or false-positive error in terms of financial and material loss. Various fault prediction approaches have been developed in the literature [11] to achieve low classification error rates. Other Support Vector Machine (SVM)-based approaches for minimizing misclassification situations have been proposed [12]. These approaches require a significant amount of training data, that is, data that contain classified information. This step is still challenging and time-consuming, particularly when working on new software programs that lack previous defect data.

Furthermore, time is a critical issue because these received much attention in the post-evaluation of the classification stage. Achieving a low classification error rate with the lowest possible risks is essential. Even more model re-training must be conducted after the prediction of the misclassification impact based on reference data, as demonstrated by Xiong, Y. and Zuo and Xiong in their paper [13], where they investigate the effect of misclassification errors to train a cost-effective neural network with different expense proportions. This method is time-consuming and requires an online training variable change despite the positive findings. There are also additional types of studies interested in evaluating sensor data online. In their study [14], Song and Deng use proof theory and intuitionistic fuzzy to continuously assess the trustworthiness of sensor data. As a result, the system can assign a fair dependability factor to sensors that give contradictory data. On the other hand, the proposed system design is extremely sophisticated and necessitates a good mathematics background.

Combining data from several sources yields significant benefits for decision-making and framework management, primarily in terms of increased trust and better-resolved system information. However, deciding on appropriate sensors to integrate for a specific task is difficult. Time and money are generally spent on respectful considerations. Instead, we present a clever AI-based solution to advanced sensor fusion that determines multiple sensor data streams based on the individual requirements, situations, and tasks. We recommend using AI automation as a pre-evaluation method in particular. This AI-based sensor assessment and clever fusion using interpretable frameworks technique can be easily applied to a wide range of sensor fusion systems.

As a result, the model's interoperability allows the candidate to follow the decision-making activity. In

addition, the suggested method can deny a request if the model is unsure about a decision. We demonstrate how to create a bi-functional system that incorporates both aspects. We concentrate on both the dynamic and static pre-evaluation of the system. A dynamic pre-evaluation evaluates sensor inputs during the classifier training phase. In contrast, a static pre-evaluation is done offline after the classifiers have been taught but before they are used in manufacturing lines. The developed system enables the evaluation of each sensor in terms of its data contribution to a predetermined categorization assignment and the hardiness of this information based on various external conditions.

Artificial intelligence (AI) has been hailed as a cure for a slew of problems in various industries [15]. It can upend old business models by opening new ones [16–18]. AI applications in the industrial business promise unique services in addition to efficiency improvements [19]. Enhancing goods with data-driven solutions is a crucial revenue generator in marketplaces with shrinking margins. It allows companies to stand out from competitors, especially in environments with ubiquitous nondata-driven services [20]. As discussed in this research [21], AI technologies are particularly valuable for commercial comprehensive supplier marketing strategies. Full-service providers (FSPs) retain ownership of intellectual products (e.g., industrial systems) and offer their use as a service in the manufacturing sector. FSP customers gain from converting procurement expenses into usage- or time-based costs, as well as the elimination of operational costs and the transfer of property control risks to the FSP. The FSP, on the other hand, gains from improved client loyalty [22] and additional revenue streams by embracing innovative payment formats. AI applications have the potential to boost the profitability of FSP marketing strategies by lowering maintenance costs and improving the availability of products or durability [23].

On the other hand, most AI applications use statistical approaches to training as part based on information [24]. These predictions allow for categorization that helps with various industrial applications and services. Cost-effective predictive maintenance (PM) and computerized predictive quality assurance (PQ) are two examples of such applications. As a result, the categorization algorithms that FSPs use must add value to consumers–i.e., service recipients (SRs)– while still being profitable.

Although most machine learning techniques are designed to turn even complicated cognitive issues into a binary classification [25], we look into classification techniques. Furthermore, the statistical structure of today's AI applications renders classifications inaccurate; nonetheless, studies show that up to 30% of decision-makers lack a basic knowledge of AI. Firms wanting to reinvent data-driven solutions based on categorization techniques must account for this imperfection in the construction of service-level agreements (SLAs) to obtain the promised benefits. As a result of the poor predictive power (PP), misclassifications resulted in breakdowns and decreased service levels. The FSP, for example, may be required to pay the SR for a lesser service level, resulting in additional costs.

Similarly, fluctuating service standards could impact an FSP's revenue. The revenue fluctuates regarding the payment system according to the PP-dependent quality of service. Low PP has little effect on a subscriber payment schedule, which provides FSPs with consistent revenue. On the other hand, a high PP allows FSPs to grow income by increasing service levels in utilization payment models. On the other hand, Low PP reduces the level of services provided and, consequently, income. The interaction of PP and different payment arrangements might positively or negatively impact the FSP's estimated net present value (NPV) [26]. As a result, FSPs must use an economic calculus to weigh the risks and benefits of using classification algorithms to pick payment arrangements.

1.1. Contribution. In Industrial automation applications, integrating blockchain, artificial intelligence, and big data constitutes the core technologies that allow dynamic data transmission. Moreover, integrating these technologies provides many features in addressing the challenges related to security, such as transparency, privacy, ensuring ownership rights, decentralization, and so on [27]. The integration of blockchain and artificial intelligence, on the other side, is still being investigated. More research studies have recommended artificial intelligence adoption using a simplified distributed system, with a focus on decentralized authentication. These research studies have failed to develop an artificial intelligence-based big data security model. Furthermore, the blockchain is not employed for big data analytics to overcome the risk of handing dynamic data into the system [28]. Instead, researchers implemented a blockchain model incorporated with distributed ledger for secure transaction processes in the industries.

In this manuscript, we introduce an improved DPoS-based consensus algorithm to increase the data transaction speed, decentralized control, and data security for IIoT networks. The novelty of developed consensus mechanism for industrial applications is as follows:

(i) To resolve the centralized security problems of IIoT, we proposed an improved DPoS consensus algorithm based on honor delegates for real-time applications.

(ii) To accelerate IoT block creation, nodes use an improved DPoS to reach a consensus for selecting delegates and store block information in the trading node.

(iii) Due to the demanding needs of enabling technologies in industrial applications, the data transmission and energy consumption are challenging tasks. To overcome these issues, we introduce delegates and honor delegate nodes–based consensus algorithm for AI-enabled IIoT.

2. Related Work

In this section, we describe the basic mechanism of artificial intelligence, big data, and blockchain for industrial application and how consensus mechanism-based AI change

industrial IoT. Blockchain, data science, and AI are the enabling technologies for industrial applications. Blockchain is mainly focused on the distributed ledge and decentralized framework for real-time applications. At the same time, data science is used for providing dynamic information and AI is employed for analyzing and classifying the sensor data in IoT applications. These innovative techniques are allowing machines to make decision and provide intelligent services. Figure 2 shows the IoT node interconnection for industrial applications.

2.1. Artificial Intelligence and Big Data Overview. Big data has been one of the most prominent research topics in recent years. Because of its huge volume, rapid velocity, and heterogeneous diversity, it differs from regular data. These traits of volume, velocity, and variety are known as the 3 Vs of big data. Later, the list was expanded to include two more Vs: value and veracity. As a result, all data that are of substantial quantity (volume), generated at a high rate (velocity), and diverse in nature (organized, semi-organized, or unorganized) are referred to as big data (variety). The value of big data analytics is that it incorporates the fourth V (value) into its qualities, making it a valuable asset to the company.

Big data analytics is a technique for analyzing large amounts of data and turning it into useful information by employing cutting-edge statistical, analytical, logistic, or artificial intelligence methods. The 3 Vs of big data, on the other hand, introduce a new set of obstacles, such as collecting, storing, exchanging, organizing, processing, analyzing, and visualizing such large amounts of data at rapid speeds [29]. Various frameworks have been built to manage large data for successful analytics in various applications for this purpose.

The digital reproduction of three primary cognitive abilities: training, thinking, and self-correction, is known as artificial intelligence (AI). Digital learning is a set of principles applied as a predictive algorithm that transforms real-world historical data into useful information [30]. The purpose of digital reasoning is to select the best rules for achieving a specific goal. Digital self-correction, on the other hand, is the continuous process of accepting the results of reinforcement learning. This approach is followed by every AI model in order to create a smart technology that can accomplish a task that would ordinarily consume a lot of time.

Machine learning, deep learning, data analysis, and principle techniques are used in the majority of intelligent systems, while reasoning and experience and understanding methods are used in others [31]. Machine learning and deep learning are two AI methods that are commonly employed nowadays. The distinctions among artificial intelligence, machine learning, and deep learning methods are frequently misunderstood.

Machine learning is a type of artificial intelligence (AI) that looks for certain trends in past data to help with decision-making. The more data we gather, the more precise is the learning process (eliminate the term big data). Machine learning can be classified into three types based on the
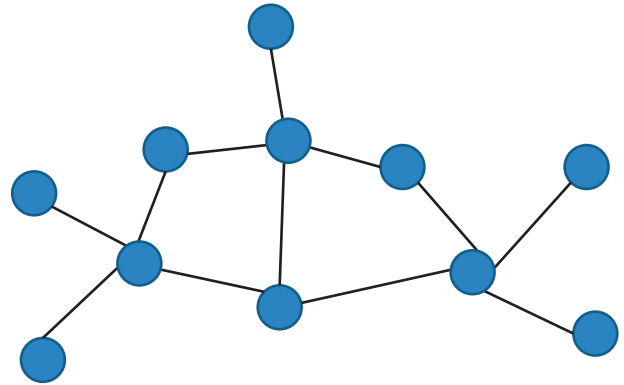
Figure 2: Graph-based IoT node generation for industrial applications.

decision process. First, supervised learning, wherein sets of data containing labelled outputs are accepted in required amounts to practice a model for categorization or future projections. Second, unsupervised learning is a type of machine learning that works with unstructured sets of data which are used for clustering and sorting. Finally, reinforcement learning collects data recordings with no labels but delivers response to the intelligent agent once specified actions take place. Linear regression, decision tree, and SVMs are the examples of supervised machine learning algorithms [32]. *K*-means and hierarchical cluster analysis fall under the unsupervised learning [33]. Lastly, Monte Carlo learning and Q-learning comes in the categories of reinforcement learning techniques [34]. Deep learning is a data mining technique inspired by the biological neural network and utilising one or more hidden units of artificial neurons. The historical data are handled repeatedly by several layers during the learning process, creating links and continuously weighting the neuron inputs for best results.

2.2. Relationship between Artificial Intelligence, Big Data, and IIoT. Real-time surveillance of physical equipment, indoor asset management, and outdoor asset management are just a few of the novel opportunities enabled by new smart sensors and IoT deployments in industrial ecosystems [35]. By integrating the physical environment to its virtual picture, IoT devices promote the real-time data gathering required for the production of a digital model of the physical component and permit the enhancement and servicing of the physical component (using smart devices). Because the IoT data indicated above is large in size, big data analytics can be useful in the building of an effective technology.

The reason for this is that industrial activities are extremely complicated, making early detection of possible issues difficult using conventional methods. Such issues, on the other side, may be easily retrieved from collected data, bringing productivity and expertise to industrial applications. However, in the industrial and technological realms, handling this massive volume of data necessitates complex approaches, structures, platforms, technologies, and algorithms. In a digital twin setting, for example, Zhang et al. [36] suggested a big data analytic system for smart account

auditing and maintenance. The relationship between AI, big data, and IIoT is depicted in Figure 3.

Cloud technology is frequently the perfect platform for processing and analyzing large amounts of data [37]. Furthermore, only by using AI technologies on the obtained data how would an intelligent digital system be created. In a nutshell, the IoT is used to collect large amounts of data from the physical world. The data are then placed into an AI model to create a digital twin. The developed digital system can then be used to improve other industrial processes.

2.3. Research Challenges and Security Issues.

The growing acceptance and accessibility of blockchain, as well as the use of IoT, data science, and AI innovations, has broadened the research problems of blockchain. These difficulties are divided into four categories.

2.3.1. Data Collection.

Data collecting from a physical device, data combining, and data exchanging with the associated blockchain are all made easier by the IoT. This procedure has the potential to be rather costly. It is possible that the digital ledger will be more expensive than the asset itself, in which case it will not be worthwhile to build the digital system. The acquired data, on the other hand, is huge, fragmented, unorganized, and noisy. As a result, more data processing is necessary to guarantee its optimal usage.

We need to use data cleaning procedures, as well as organize, rearrange, and homogenize the data. Furthermore, keeping such a massive volume of data under control is a huge task. Furthermore, the fundamental machine learning techniques require a specific quantity of data for training reasons in order to enhance the reliability of the blockchain model.

2.3.2. Challenges in Big Data.

The rapid use of IoT technology in the industries has resulted in massive volumes of monitoring (sensor) data being generated. To this goal, improved infrastructures, foundations, platforms, techniques, and strategies are needed to represent, preserve, share, analyze, and evaluate the raw data in big data and analytics. Edge and cloud services platforms could also be used to handle digital twin related data. Edge computing, in particular, allows for dispersed computation at the network's edge, with collective analysis taking place in the cloud. However, data processing on the cloud may result in a longer response time.

2.3.3. Analysis of Raw Data.

As described in the literature, artificial intelligence–based techniques for big data played an important role in industry for decision-making. However, choosing a certain model from hundreds of machine learning with unique settings is difficult. To various applications and data sources, each intelligence has various levels of accuracy and efficiency. On the other hand, accuracy might have a negative impact on efficiency. As a result, choosing the proper optimization algorithm and functionalities is difficult depending on the motivation and
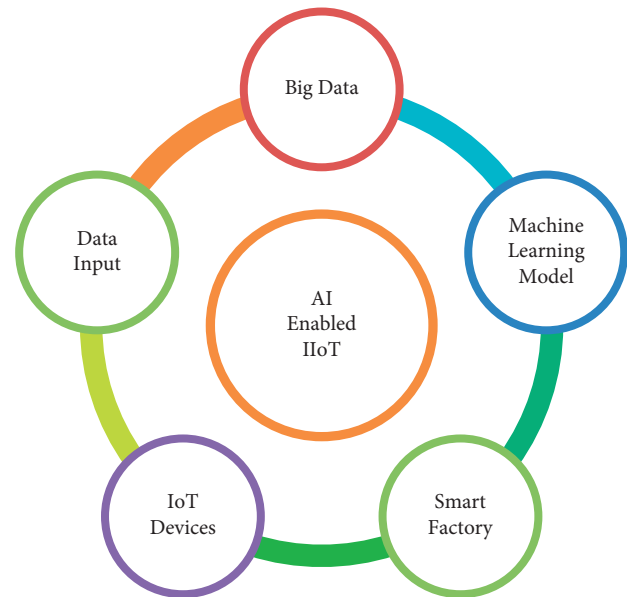
FIGURE 3: AI enabled IIoT network integration.

implementation of industry automation. Furthermore, there are less realistic deployments of intelligence for industry 4.0 revolution in the literature, which adds to the difficulties.

2.3.4. Challenges in Privacy and Security.

Some manufacturing sectors, such as sensor data, product-related information, and human management ledger are deemed sensitive and may demand strict security and privacy guarantees. First, because IoT devices are involved in digital twinning, the privacy of the fundamental communication systems must be prioritized. Furthermore, the enormous amount of asset-related data must be securely held to avoid data theft from both inside and outside attacks.

2.4. Need of Blockchain Technology in IIoT.

In recent times, there has been a lot of studies into the privacy and security of interaction among IoT devices. Blockchain technology is a new use in IIoT networks, and its effective deployment has been the focus of much research. The IIoT benefits from the blockchain's decentralization, data integrity, cryptography privacy, fault tolerance, data security and identification, and consensus mechanism [38]. Several research studies compared popular blockchain platforms, including cryptocurrency, Ethereum [39], Hyperledger-Fabric [40], and IOTA [41], and discussed the advancement of smart contracts and its practicability in the industry, IOTA offers free transactions designed specifically for device to device communication, but it lacks the maturity of Ethereum and Public blockchain.

Blockchain technology has clearly evolved in people's perception as scientific research and innovations have progressed, and it has become a topic of studies by scholars and researchers. Industry and academics are paying growing attention to it. People have recognised the one-of-a-kind extraordinary development that distributed ledger may

bring about and have committed in the development on business elements such as banking, healthcare, and traceability. The distributed system is speeding up the maturity and industrial adoption of blockchain technology. Presently, China is creating its own blockchain technology; the competitive market structure and the separate copyright system are being developed [42].

At the same time, relevant techniques and developed ecology are integrating sectors, such as energy, healthcare, and agriculture. The decentralized security is a new kind of innovative platform in that digital information such as random data blocks are used to authenticate user and provide the data privacy through consensus algorithm. The blockchain technology is implemented to ensure the security and privacy of data transmission between nodes.

2.5. Consensus Algorithm for Industrial Applications. A consensus algorithm is a collection of rules governing how a decentralized network is supposed to work. These principles outline the basic roles of various parts, how they interact, and the criteria that must be met in order for them to function correctly. A consensus algorithm specifies the rules that must be followed in order to establish an agreement, as well as the procedures that should be done under what situations. The proof of work (POW) technique states that as long as such a node could generate a block which adheres to the desired value, the entire network can verify it [43]. In a distributed system, a consensus algorithm is a technique for resolving data synchronisation between nodes that do not trust each other.

PoX (proof-of-X) decision techniques for blockchain systems without authorities have recently emerged and developed, with all techniques focusing on network transactions [43]. However, because there is no agreement, transaction verification is delayed, which is incompatible with most dynamic IIoT devices that demand real-time validation. The Equihash method [45] is a proof-of-work (PoW) agreement technique based on the generalised birthday dilemma in which a fundamental cryptography implementation is difficult. This is a memory-dependent consensus technique that sets the burden based on the nodes' storage sizes. It requires a lot of storage to provide evidence, but it can achieve quick confirmation. Although this design enhances the cost-effectiveness of ASIC devices, the application's security has yet to be validated.

The Ouroboros techniques deployed in Ref. [46] is a distributed system-based proof of stake (PoS) consensus model. The techniques develop a tight security guarantee consensus procedure and push the PoS consensus process via a reward system. This reward is to confirm that non-malicious devices maintain a nash equilibrium and also prevent security breaches affected by selfish block creation. Da Xu and Viriyasitavat [47] combined the PoW and PoS consensus concepts for security transactions. The PoW method is utilised for the acquisition of tickets in the early stages of the process. When the blockchain system has acquired sufficient assets, the PoS algorithm is utilised to ensure the network's long-term safety. The PoS algorithm

provides direct correlation between coin age and time that is converted to an exponentially decaying rate. This approach helps the rate of growth of coin age to reach zero over the period of time and prohibits the accumulation of money. However, the approach increases computation time and necessitates a large amount of network memory size.

Based on the original Paxos algorithmic concept, Moraru et al. [48] created EPaxos consensus method. Creating the dependency, accepting the request, and completing the phase are the three stages of the method. Each proposition has characteristics, such as gathering and pattern numbering, in addition to the intrinsic data. To establish the execution order of competing proposals, the ideas of a quick stream, slow stream, and dependency graph are presented. Only implementation situations with few or no conflicts are acceptable for the technique.

For the distributed ledgers, Sousa et al. [49] introduced the Byzantine fault-tolerant consensus protocol. To model the continuity of Byzantium fault-tolerance, this protocol employs a probability remuneration network. Although the protocol offers some benefits in terms of transmission capacity and faster transaction time, the execution flow could be improved. On the basis of credit, Yeow et al. [50] suggested an enhanced practical Byzantine fault-tolerant (PBFT) consensus protocol. The consensus protocol was enhanced, credit assessment predicated on a coalition chain was created, and the system was brought into a feedback loop by including a lightweight integrity method. The checkpoints protocol was changed to allow devices to enter and depart the network on demand, increasing the platform's adaptability.

Lin et al. [51] developed a ring signature-based modified PBFT consensus technique. The PBFT technique, the ElGamal cryptographic signature encryption method, and the ring signature concept were all presented. The efficiency and secrecy of a ring signature technique based on the ElGamal technique were then investigated. The ring signature strategy was optimized to enhance the PBFT computation signature and validation procedure, allowing nodes to enter and depart the network continuously. The suggested solution surpassed the original PBFT computation fault-tolerant percentage.

Delegated Proof of Stake (DPoS) was created by Larimer and deployed initially for the BitShares project [52]. The DPoS consensus process is separated into two parts: the first is the election of witnesses (block creators), and the second is the generation of blocks. Witnesses are simply authorized for confirming the transactions, validating the signature, and timestamping it; they are not allowed to trade. They each produce one block every three seconds, and if a witness fails to perform the task within the time limit, it is ignored and substituted by the next one. Each network node has the ability to vote for its own dedicated witness, and the more smart contract stakes he or she has, the more likely he or she is to be a witness. However, because of the core method whereby each witness node generates blocks in succession, the identification of the witness is already established and stable, making the distributed ledger network more open to fraud assaults.

Furthermore, achieving fairness by solely employing DPoS is challenging, as it will only allow those with more resources to become voters. In addition, while employing the PoW alone, the block period is around 10 minutes, wasting a significant amount of computer and energy resources. For this third issue, this paper proposes an honor mechanism akin to that used in reputation to select the delegate node that causes the consensus to be harmed and replace them.

There are two types of nodes in the consensus algorithm, according to our findings. The delegate node, for example, is a node that creates or validates transactions and contributes in the consensus mechanism. The honor delegate node, on the other hand, serves as a contender for the delegate node and is used to replace it if it fails. These two types of nodes are termed as delegate nodes for conforming consensus mechanism.

2.6. Proposed Blockchain-Based Consensus Mechanism for AI-Enabled IIoT. To resolve the security and privacy issues of AI and enable IIoT, considering the distributed ledge of sensor information, we proposed an improved DPoS consensus mechanism based on honor voting system for industrial application. We construct a consensus mechanism–based ledger for sensor data storage in IoT device system solutions for smart industrial automation because of the transparency and data integrity of blockchain. Data cannot be changed by distant attackers wanting to quickly get into the device for harmful modification. Since IoT sensor data, such as identification, password protection, application settings, and behavioral records, may be safely kept in distributed ledger, in this paper, we employed an improved DPoS algorithm to create a consensus for producing blocks including sensor information, which speeds up block generation.

2.7. Improved DPoS Consensus Algorithm. The most appropriate delegates cannot be picked for block formation due to the inaccuracy of voting choices and the inaccuracy of vote calculation. An improved DPoS algorithm is introduced in this manuscript to be more effective, versatile, and precise in choosing suitable delegates. Improved DPoS algorithm is made up of three parts. The first is an honor voting system, which yields a collection of delegates phrases for each voting node. The second step is to create an improved voting function that will be used to determine each node's value. The highest number of honor voting, the better the node's chances of becoming a delegate. To complete the voting process, the final step is to determine the divergence level.

2.8. The Basic Concept of Improved DPoS Algorithm. According to the literature, the major drawbacks of the public blockchain consensus mechanism are that the distributed ledge techniques are more permissioned and insecure, and its difficulty in creating blocks is considered as a critical limitation to our technological needs. To overcome these issues, the DPoS-based consensus algorithm can solve the security problems because the DPoS algorithm can

greatly enhance the authentication and also reduce energy consumption. On the other side, the PoW mechanism greatly reduces the energy consumption because every node has the right to create blocks.

However, achieving fairness by solely employing DPoS is challenging, as it will only allow those with more money to become voters. Furthermore, while employing the PoW alone, the block interval is around 10 minutes, wasting a significant amount of computing power sources. For the power consumption issue, this paper proposes an honor voting mechanism akin to that used in the modern system to reduce the latency of the voting node that causes the consensus to be harmed and replace them. We have considered that there are two types of nodes in the consensus algorithm. The one is the voting node, which creates blocks and takes part in the authentication process. The second node is termed as a honor voting node, and it has special voting privilege. With the special voting privilege, the honor voting node can replace the node when it fails to perform. And, these kinds of proposed nodes are generally considered as the consensus nodes.

2.9. The Development of Consensus Mechanism. We separate the consensus procedure into two parts. To begin with, utilise the PoW concept to select a set number of suitable nodes from the entire network, and employ stake voting to select consensus nodes. In this network, the top 101 nodes serve as delegate nodes and the remainder nodes serve as honor delegates. The delegate nodes then record the transactions in a block and disseminate it to all consensus nodes for consensus in the second phase. The block will be added to the blockchain if it is successfully verified by more than half of the consensus nodes. The process steps of an honor delegate node selection based on improved DPoS algorithm is shown in Figure 4.

As immediately as the malicious node is discovered, we implement an honor voting method to regenerate it. When a delegate node is discovered to be malicious, it is added to the honor delegates' nodes set, and the rank of all existing delegate nodes is reduced by one. The node in the set of honor delegate nodes with sequence number is moved to the delegate nodes set and ranked last in the delegate nodes set. The malicious node is sorted at the end of the collection of honor delegate nodes, while the identifiers of all the remaining honor nodes are decremented by one. Algorithm 1 describes the proposed mechanism for honor node selection in the IoT network.

The N number of nodes with the maximum votes can be selected as delegate nodes in improved DPoS, and the shortlist of delegates nodes will be updated every 20 hours, just like in DPoS. The delegate node's reputation and witness identification will be revoked if someone is discovered to have a poor rate of creating blocks or to be engaging in harmful activities.

The module for choosing consensus networks: all nodes in the public ledger are successfully prepared in order to allocate various tasks to different types of nodes, which are primarily separated into consensus nodes (such as delegate
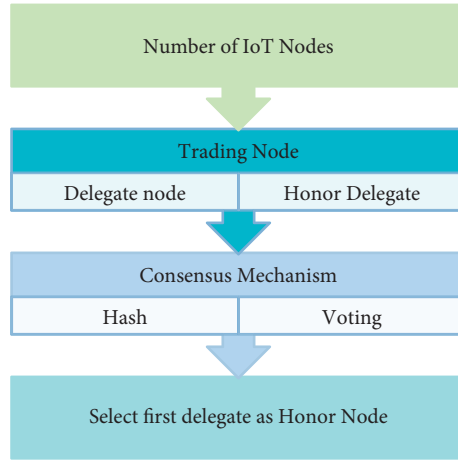
FIGURE 4: Selecting an honor delegate node in the Improved DPoS Algorithm.

The algorithm for selection of honor delegate node
Input: Voting results among N nodes, nonce
Output: honor delegates
Transmit (nonce, N)
N_i: Hash (hash (BlockHead), nonce)
while (Hash (hash (BlockHead), nonce) > delegates)
Calculate the vote of each delegate and sort them
if number of vote of first delegate ≥ second delegate
Select the first delegate as the honor delegate
If the votes of different delegates are same
Calculate the vote deviation percentage
Select small deviation percentage delegate as honor node
end if
end if

ALGORITHM 1: Honor delegates node selection algorithm.

nodes and honor nodes) and transaction network. Trading nodes are responsible for generating transactions, while consensus nodes are responsible for generating and verifying blocks. Figure 5 shows the delegate node selection using the graph method. The consensus module is responsible for executing the entire process from block creation to block confirmation. The module for degrading malicious nodes is as follows: When a malicious node is discovered, the improved DPoS algorithm switches to a module that replaces honor node and sorts the delegate nodes.

The decentralized network ecosystem is described as a peer-to-peer network made up of all branches in the architecture, with consensus nodes and trade nodes being the two types of nodes in this network. The consensus network is a sub-network ecosystem made up of delegate nodes that will change when the voting number in the improved DPoS algorithm updates. Conversely, the trading sub-network is the network architecture comprising of trading nodes which is not static. The trading nodes update them after each round of selecting the delegate node.

Only during the time between the formation of the current delegate nodes and the beginning of another round of honor node selection will the trading networks be secure.

As previously stated, the nodes that participate in the consensus process rather than generating transactions are divided into two categories: delegate node and honor nodes. The trade nodes are accountable for the production, transmitting, and storage of distributed ledger, while the witness nodes take turns recording trades into a block and transmitting it to the other delegate nodes for validation.

3. Result Evaluation and Discussion

In this section, we describe the performance assessment of the blockchain-based consensus mechanism for AI-enabled IIoT network. We present the comparison result of the proposed work with existing mechanism such as PoS, PoW, and DPoS in terms of important parameters which includes block creation approach, block generation time, and energy resources. To carry out security performance, we include existing consensus mechanism. Each work represents the artificial intelligence–based big data analysis at the IoT device with decentralized control. The proposed blockchain-based consensus mechanism provides distributed architecture to enhance security and privacy at the IoT device.
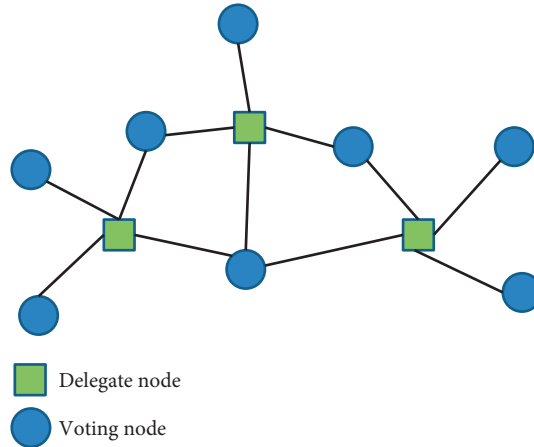
FIGURE 5: Graph-based delegate nodes selection in consensus mechanism.

TABLE 1: Performance of various consensus algorithm.

Consensus mechanism	Proof-of-work (PoW)	Proof-of-stack (PoS)	Improved delegate PoS (DPoS)
Mechanism for block generation	Computing power	Stake	Stake votes
Security issues	Constant power	Inactive nodes	Malicious nodes
Energy consumption	Very high	High	Low
Average block generation time	10 min	65 sec	5 sec
Reliability	High	Low	Low
Robustness	High	High	High
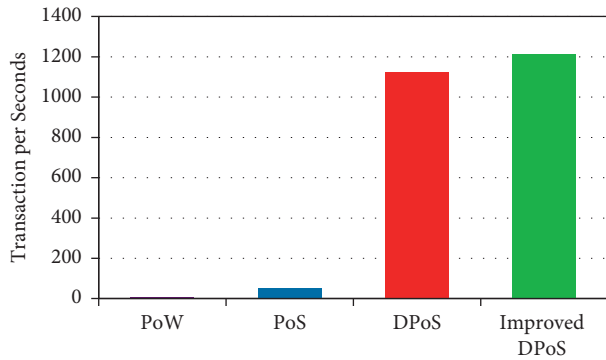
FIGURE 6: Transaction data rate of various consensus mechanisms.

The distributed consensus mechanism–based AI-enabled IoT architecture resolved the real-time security issues and also reduced the energy consumption. An improved DPoS algorithm suggested AI-based decentralized IoT network for big data analysis in real-time, and it will overcome the issue of data storage. The proposed mechanism introduces the trading node in which the actual block information is stored. The implementation of big data analytic is developed on the IIoT blocks to evaluate the scalability and robustness of improved DPoS. We presented the performance of the proposed method with existing consensus algorithm research shown in Table 1. According to the security metric, the proposed algorithm provides better results compared with PoS and PoW in terms of energy consumption.
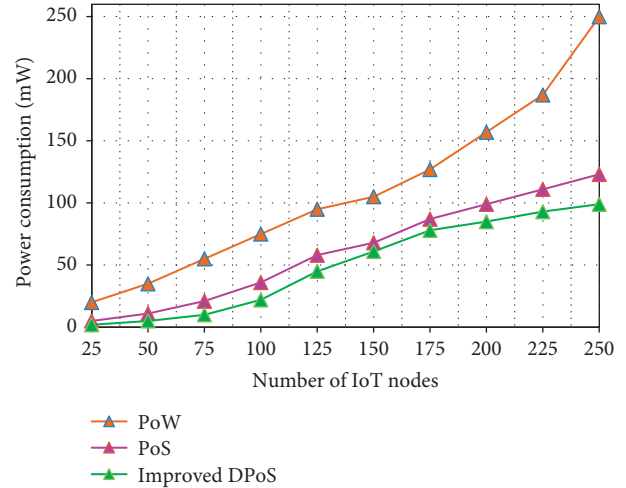
FIGURE 7: Energy consumption comparison of various consensus mechanisms.

To analyze the data transaction rate of improved DPoS algorithm, we investigate the performance with existing methods such as a PoW, PoS, and DPoS mechanism [43]. The transaction rate of PoW is very low because of its computation time. In PoW, blocks are verified based on the computing power. In the same way, in the case of PoS the blocks are verified through stake methods and it will require more transaction time. In order to reduce the transaction time, DPoS is proposed based on the stake voting mechanism.

Likewise, our proposed algorithm verified through honor delegates which required less time for verification. Therefore, compared with other three mechanisms, our improved DPoS increases the transaction rate.

Figure 6 clearly shows that our improved DPoS algorithm has more TPS than other mechanisms. To reduce the energy resource of the blockchain-based consensus mechanism, we developed an improved DPoS algorithm for IIoT devices. In order to overcome the computing resources problems in the PoW and PoS mechanism of decentralized ledger-based industrialized IoT devices, we propose to combine artificial intelligence and blockchain technology. In the decentralized improved DPoS consensus algorithm implemented to enhance the data privacy and reduce energy resources for big data analysis, the main reason to introduce delegates and honor nodes-based consensus mechanism for IIoT devices is to reduce overall energy consumption, as it will verify the blocks in the idea of stake voting mechanism and also replace malicious nodes through honor delegates. To develop smart contracts between nodes, we used DPoS-based stake voting mechanism, which is denoted as delegates nodes.

Figure 7 represents the energy consumption analysis of the proposed algorithm and other two existing models. In the PoW mechanism, if the number of blocks increases, the energy consumption also increases because in this algorithm the blocks are created based on the computing power of the particular system. Therefore, the PoW model required more resources than other models. In the case of the proposed algorithm, the blocks are created through stake vote which will be part of the IoT network, as the energy consumption is very less compared to the PoW model.

Finally, the different performance evaluations of the developed blockchain-based AI-enabled IIoT network conforms that an improved DPoS consensus algorithm increases data transaction rate per seconds and reduces the energy resources. Overall, the proposed consensus mechanism is most applicable for AI-based IIoT applications in order to analyze data in a secure manner with less energy resources.

4. Conclusion

A combined blockchain and artificial intelligence–based consensus algorithm for big data analysis in IoT applications are introduced in this manuscript. This work aims to develop efficient and reliable IoT data transactions at the industrial level. The suggested DPoS consensus algorithm performance was evaluated using security and energy consumption metrics. An improved DPoS was implemented to blockchain-based AI for decentralized control in IIoT. The experimental analysis is presented to evaluate the performance of the suggested consensus mechanism AI-enabled IIoT applications with distributed and secure big data analytics. In terms of reliability, speed, privacy, and security, the experimental results show the efficiency of the proposed algorithm compared with existing mechanisms. According to the TPS results, the integration of blockchain with artificial intelligence successfully addresses the issues of getting

high accuracy, security, and low latency through a decentralized network. The proposed consensus algorithm successfully overcomes the difficulties of accuracy, latency, and security by combining blockchain and artificial intelligence and also addresses the energy consumption issue [44].

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] "1. UN General Assembly, Transforming Our World: The 2030 Agenda for Sustainable Development," 2015, https://www.refworld.org/docid/57b6e3e44.html.

[2] Q. Shi, Z. Zhang, Y. Yang, and X. B. C. Shan, "Artificial intelligence of things (AIoT) enabled floor monitoring system for smart home applications," *ACS Nano*, vol. 15, no. 11, pp. 18312–18326, 2021.

[3] W. Bronner, H. Gebauer, C. Lamprecht, and F. Wortmann, "Sustainable AIoT: how artificial intelligence and the internet of things affect profit, people, and planet," *Connected Business*, vol. 1, pp. 137–154, 2021.

[4] G. Rong, Y. Xu, X. Tong, and H. Fan, "An edge-cloud collaborative computing platform for building AIoT applications efficiently," *Journal of Cloud Computing*, vol. 10, no. 1, p. 36, 2021.

[5] G. Premsankar, M. Di Francesco, and T. Taleb, "Edge computing for the internet of things: a case study," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1275–1284, 2018.

[6] R. S. Peres, X. Jia, J. Lee, and K. A. W. J. Sun, "Industrial artificial intelligence in industry 4.0 - systematic review, challenges and outlook," *IEEE Access*, vol. 8, pp. 220121–220139, 2020.

[7] H. Ramamurthy, B. S. Prabhu, R. Gadh, and A. M. Madni, "Wireless industrial monitoring and control using a smart sensor platform," *IEEE Sensors Journal*, vol. 7, no. 5, pp. 611–618, 2007.

[8] K. S. Ladha and M. Eikermann, "Codifying healthcare - big data and the issue of misclassification," *BMC Anesthesiology*, vol. 15, no. 1, p. 179, 2015.

[9] Y. Xiong, Y. Ge, L. Grimstad, and P. J. From, "An autonomous strawberry-harvesting robot: design, development, integration, and field evaluation," *Journal of Field Robotics*, vol. 37, no. 2, pp. 202–224, 2020.

[10] T. Wossen, A. Alene, T. Abdoulaye, and S. I. Y. V. Feleke, "Poverty reduction effects of agricultural technology adoption: the case of improved cassava varieties in Nigeria," *Journal of Agricultural Economics*, vol. 70, no. 2, pp. 392–407, 2019.

[11] X. Xu, T. Chen, and M. Minami, "Intelligent fault prediction system based on internet of things," *Computers & Mathematics with Applications*, vol. 64, no. 5, pp. 833–839, 2012.

[12] Y. Lin, Y. Lee, and G. Wahba, "Support vector machines for classification in nonstandard situations," *Machine Learning*, vol. 46, no. 1/3, pp. 191–202, 2002.

[13] R. Zuo and Y. Xiong, "Big data analytics of identifying geochemical anomalies supported by machine learning methods," *Natural Resources Research*, vol. 27, no. 1, pp. 5–13, 2018.

[14] Y. Song and Y. Deng, "A new method to measure the divergence in evidential sensor data fusion," *International Journal of Distributed Sensor Networks*, vol. 15, no. 4, Article ID 155014771984129, 2019.

[15] O. Müller, I. Junglas, J. v. Brocke, and S. Debortoli, "Utilizing big data analytics for information systems research: challenges, promises and guidelines," *European Journal of Information Systems*, vol. 25, no. 4, pp. 289–302, 2016.

[16] C. Chen, "Business intelligence and analytics: from big data to big impact," *MIS Quarterly*, vol. 36, no. 4, p. 1165, 2012.

[17] J. Weking, M. Stöcker, M. Kowalkiewicz, M. Bohm, and H. Krcmar, "Archetypes for industry 4.0 business model innovations," in *Proceedings of the 24th Americas Conference on Information Systems*, New Orleans, LA, USA, October 2018.

[18] T.-M. Choi, S. W. Wallace, and Y. Wang, "Big data analytics in operations management," *Production and Operations Management*, vol. 27, no. 10, pp. 1868–1883, 2018.

[19] F. Kache and S. Seuring, "Challenges and Opportunities of Digital Information at the Intersection of Big Data Analytics and Supply Chain Management," *International journal of operations & production management*, vol. 37, no. 1, 2017.

[20] R. X. R. Huber, L. C. Püschel, and M. Röglinger, "Capturing smart service systems: Development of a domain-specific modelling language," *Information Systems Journal*, vol. 29, no. 6, pp. 1207–1255, 2020.

[21] L. Fabri, B. Häckel, A. M. Oberländer, J. Töppel, and P. Zanker, "Economic perspective on algorithm selection for predictive maintenance," in *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm & Uppsala, Sweden, June 8-14, 2019.

[22] J. A. Guajardo, M. A. Cohen, S. H. Kim, and S. Netessine, "Impact of performance-based contracting on product reliability: an empirical analysis," *Management Science*, vol. 58, no. 5, pp. 961–979, 2012.

[23] R. Clarke, "Big data, big risks," *Information Systems Journal*, vol. 26, no. 1, pp. 77–90, 2016.

[24] D. Lazer, R. Kennedy, G. King, and A. Vespignani, "The parable of Google Flu: traps in big data analysis," *Science*, vol. 343, no. 6176, pp. 1203–1205, 2014.

[25] Z. Zhang, J. T. Hummel, J. Nandhakumar, and L. Waardenburg, "Addressing the key challenges of developing machine learning AI systems for knowledge-intensive work," *MIS Quarterly Executive*, vol. 19, no. 4, pp. 221–238, 2021.

[26] D. Halbheer, D. L. Gärtner, E. Gerstner, and O. Koenigsberg, "Optimizing service failure and damage control," *International Journal of Research in Marketing*, vol. 35, no. 1, pp. 100–115, 2018.

[27] M. Javaid, A. Haleem, R. Pratap Singh, S. Khan, and R. Suman, "Blockchain technology applications for Industry 4.0: a literature-based review," *Block: Research and Applications*, vol. 2, no. 4, Article ID 100027, 2021.

[28] N. Tariq, M. Asim, F. Al-Obeidat et al., "The security of big data in fog-enabled IoT applications including blockchain: a survey," *Sensors*, vol. 19, no. 8, p. 1788, 2019.

[29] J. S. Johnson, S. B. Friend, and H. S. Lee, "Big data facilitation, utilization, and monetization: exploring the 3Vs in a new product development process," *Journal of Product Innovation Management*, vol. 34, no. 5, pp. 640–658, 2017.

[30] S. Tolan, A. Pesole, F. Martínez-Plumed, E. Fernandez-Macias, J. Hernandez-Orallo, and E. Gomez, "Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks," *Journal of Artificial Intelligence Research*, vol. 71, pp. 191–236, 2021.

[31] D. M. Dimiduk, E. A. Holm, and S. R. Niezgoda, "Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering," *Integrating Materials and Manufacturing Innovation*, vol. 7, no. 3, pp. 157–172, 2018.

[32] Y. Leng, X. Xu, and G. Qi, "Combining active learning and semi-supervised learning to construct SVM classifier," *Knowledge-Based Systems*, vol. 44, pp. 121–131, 2013.

[33] A. D. Peterson, A. P. Ghosh, and R. Maitra, "Merging K-means with hierarchical clustering for identifying general-shaped groups," *Stat*, vol. 7, no. 1, p. 172, 2018.

[34] J. Fan, Z. Wang, Y. Xie, and Z. Yang, *A Theoretical Analysis of Deep Q-Learning*, pp. 486–489, Learning for Dynamics and Control, Stanford, CA, 2020.

[35] J. Astill, R. A. Dara, E. D. Fraser, B. Roberts, and S. Sharif, "Smart poultry management: smart sensors, big data, and the internet of things," *Computers and Electronics in Agriculture*, vol. 170, Article ID 105291, 2020.

[36] J. Zhang, X. Yang, and D. Appelbaum, "Toward effective big data analysis in continuous auditing," *Accounting Horizons*, vol. 29, no. 2, pp. 469–476, 2015.

[37] C. Ji, Y. Li, W. Qiu, U. Awada, and K. Li, "Big data processing in cloud computing environments," in *Proceedings of the 12th international symposium on pervasive systems, algorithms and networks*, pp. 17–23, San Marcos, TX, USA, 13-15 December 2012.

[38] Q. Wang, X. Zhu, Y. Ni, L. Gu, and H. Zhu, "Blockchain for the IoT and industrial IoT: a review," *Internet of Things*, vol. 10, Article ID 100081, 2020.

[39] E. K. Wang, Z. Liang, C. M. Chen, S. Kumari, and M. K. Khan, "PoRX: a reputation incentive scheme for blockchain consensus of IIoT," *Future Generation Computer Systems*, vol. 102, pp. 140–151, 2020.

[40] S. Figueroa-Lorenzo, J. Añorga, and S. Arrizabalaga, "Methodological performance analysis applied to a novel IIoT access control system based on permissioned blockchain," *Information Processing & Management*, vol. 58, no. 4, Article ID 102558, 2021.

[41] B. Shabandri and P. Maheshwari, "Enhancing IoT security and privacy using distributed ledgers with IOTA and the tangle," in *Proceedings of the 6th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 1069–1075, Noida, India, 07-08 March 2019.

[42] H. Hou, "The application of blockchain technology in E-government in China," in *Proceedings of the 26th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–4, Vancouver, BC, Canada, 31 July 2017 - 03 August 2017.

[43] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, "On the security and performance of proof of work blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 3–16, Vienna, Austria, October 2016.

[44] E. K. Wang, R. Sun, C. M. Chen, Z. Liang, S. Kumari, and M. Khurram Khan, "Proof of X-repute blockchain consensus protocol for IoT systems," *Computers & Security*, vol. 95, Article ID 101871, 2020.

[45] A. Biryukov and D. E. Khovratovich, "Equihash: asymmetric proof-of-work based on the generalized birthday problem," *Ledge*, vol. 2, pp. 1–30, 2017.

[46] C. Badertscher, P. Gaži, A. Kiayias, A. Russell, and V. Zikas, "Ouroboros genesis: composable proof-of-stake blockchains with dynamic availability," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 913–930, Toronto, Canada, October 2018.

[47] L. D. Xu and W. Viriyasitavat, "Application of blockchain in collaborative Internet-of-Things services," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1295–1305, 2019.

[48] I. Moraru, D. G. Andersen, and M. Kaminsky, "There is more consensus in egalitarian parliaments," in *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*, pp. 358–372, China, November 2013.

[49] J. Sousa, A. Bessani, and M. Vukolic, "A byzantine fault-tolerant ordering service for the hyperledger fabric blockchain platform," in *Proceedings of the 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 51–58, Luxembourg, Luxembourg, 25-28 June 2018.

[50] K. Yeow, A. Gani, R. W. Ahmad, J. J. P. C. Rodrigues, and K. Ko, "Decentralized consensus for edge-centric internet of things: a review, taxonomy, and research issues," *IEEE Access*, vol. 6, pp. 1513–1524, 2018.

[51] C. Lin, D. He, X. Huang, X. Xie, and K. K. R. Choo, "PPChain: a privacy-preserving permissioned blockchain architecture for cryptocurrency and other regulated applications," *IEEE Systems Journal*, vol. 15, no. 3, pp. 4367–4378, 2021.

[52] F. Schuh and D. Larimer, "Bitshares 2.0: General Overview," Error! Hyperlink reference not valid, 2017.

Hindawi

Research Article

Deep Learning Approaches for Cyberbullying Detection and Classification on Social Media

Neelakandan S,[1] Sridevi M,[2] Saravanan Chandrasekaran,[3] Murugeswari K,[4] Aditya Kumar Singh Pundir [ID],[5] **Sridevi R,[6] and T.Bheema Lingaiah** [ID][7]

[1]*Department of CSE, R.M.K Engineering College, Kavaraipettai, India*
[2]*Department of CSE, Anurag University, Telangana, India*
[3]*Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India*
[4]*School of Computing Science and Engineering, VIT Bhopal University, Bhopal, India*
[5]*Department of ECE, Arya College of Engineering and Information Technology, Kukas, Rajasthan, India*
[6]*Department of CSE, K.Ramakrishnan College of Engineering, Samayapuram, India*
[7]*School of Biomedical Engineering, Jimma Institute of Technology, Jimma, Ethiopia*

Correspondence should be addressed to T.Bheema Lingaiah; bheema.lingaiah@ju.edu.et

As a result of the ease with which the internet and cell phones can be accessed, online social networks (OSN) and social media have seen a significant increase in popularity in recent years. Security and privacy, on the other hand, are the key concerns in online social networks and other social media platforms. On the other hand, cyberbullying (CB) is a serious problem that needs to be addressed on social media platforms. Known as cyberbullying (CB), it is defined as a repetitive, purposeful, and aggressive reaction performed by individuals through the use of information and communication technology (ICT) platforms such as social media platforms, the internet, and cell phones. It is made up of hate messages that are sent by e-mail, chat rooms, and social media platforms, which are accessed through computers and mobile phones. The detection and categorization of CB using deep learning (DL) models in social networks are, therefore, crucial in order to combat this trend. Feature subset selection with deep learning-based CB detection and categorization (FSSDL-CBDC) is a novel approach for social networks that combines deep learning with feature subset selection. The suggested FSSDL-CBDC technique consists of a number of phases, including preprocessing, feature selection, and classification, among others. Additionally, a binary coyote optimization (BCO)-based feature subset selection (BCO-FSS) technique is employed to select a subset of features that will increase classification performance by using the BCO algorithm. Additionally, the salp swarm algorithm (SSA) is used in conjunction with a deep belief network (DBN), which is known to as the SSA-DBN model, to detect and characterize cyberbullying in social media networks and other online environments. The development of the BCO-FSS and SSA-DBN models for the detection and classification of cyberbullying highlights the originality of the research. A large number of simulations were carried out to illustrate the superior classification performance of the proposed FSSDL-CBDC technique. The SSA-DBN model has exhibited superior accuracy to the other algorithms, with a 99.983 % accuracy rate. Overall, the experimental results revealed that the FSSDL-CBDC technique beats the other strategies in a number of different aspects.

1. Introduction

In recent years, people utilize virtual meeting platforms in their daily lives using global online social network (OSN) to facilitate communication. This network helps users for finding new friends and increases their connections around

the world. Furthermore, sharing of data and opinions is the significant features of OSN [1–3]. In recent years, the rate of utilizing OSN has rapidly increased. With OSNs such as Facebook, Google+, LinkedIn, Twitter, VKontakte, Mixi, and Sina Weibo, a Japanese social network turned into the desired manner of transmission for billions of everyday

active users. A user consumes maximal time for updating their content, interacting with primary user, and browsing others' account for finding particular data that are the major implication of social network website. OSN could remove the economical and geographical barriers among the users for sharing information and communication. In addition, OSN is highly beneficial to attain the objectives like amusement, education, search for jobs, etc. The popularity of OSN leads to high risks of an attack on the OSN users. Several OSN users expose their private data and that act as a proposal for the attacker to perform specific malicious activities [4, 5].

The current extensive nature of cyberbullying (CB) has enlarged the significance of its recognition. As per the survey, nearly 43% of the teenagers in the US alone have been stated to be the targets of CB at a certain point. CB is deliberated as a novel or electric method of conventional bullying. CB is determined as an aggressive, repeated, and intended response determined by an individual/group toward other individuals/groups, which is created using information and communication technology (ICT) methods like the internet, mobile phones, and social media [6]. The whole CB events are executed in internet broadcasting instead of in a physical system. The CB contains hate letters transferred by e-mails, social networking, and so on, via public/private computers/using private mobile phones. It is raised as a severe threat among the states. Current research displays the ratio as improved to be about 59% in the US. CB has a similar, when it is not a better, negative impact on the victim against conventional bullying, since the predators generally attack a victim on the aspect that an individual could not alter (viz., ethnicity, physical appearance, skin color, and religion), which leaves a deep and long-lasting effect on the victim. Occasionally, the related humiliations are sufficient for pushing the victim to self-infliction of suicide/harm.

A study in [7] displayed that suicide thoughts tend to rise among teenagers because of the disclosure of various types of CB. Although precautions are occupied, the redevelopment of victims of CB cases is challenging for society and families. Self-hate, dominance, isolation, and reaction to the socializing procedure lead to troubled and unhappy adults. Furthermore, this mental imbalance could alone make upcoming bullies. Among many problems, which create the recognition of CB in OSN is highly complicated, current advanced solution for detecting CB does not determine the possibility of bullying types in their detection method [8]. We specified the different kinds of CB, which could arise on the web, and it is not possible for assuming that a similar detection method would be effective in finding all types of bullying.

The main limitation present in the current detection system using CB study is the absence of input data. This study is traditionally executed on an available dataset/surveyed data, while the victim's/perpetrators are permitted for reporting the impression. Another issue with automatic CB identification is determining the most appropriate operation on CB material that takes into account the available studies in the CB detection region in order to achieve the goal of automatic detection accurately recognizing CB actions, which is another issue with automatic CB identification. It becomes more difficult to determine the actions as a result of this, and well-developed tools for combining the information via an autonomous decision technique are necessary [9]. To achieve the goal of automatic detection to precisely recognize CB actions, a CB detection zone was created. The automated decision-making is the process of making decisions without human intervention. Inferred data or digitally developed profiles can also be used to make these decisions. A preprogrammed algorithm and criteria may be used in an online loan decision or a recruitment aptitude examination. Administration heavily relies on automation. Automated systems can increase administrative decision-making consistency, correctness, and transparency, and enable new service delivery choices in the relevant areas and with suitable supervision. According to the findings of this study, machine learning was utilized to detect artificial CB material based on a number of psychological and common characteristics. The CB detection rate of this intelligent system has been reported to be lower, and it has been found to be mostly confined to a person writing a comment in the text [10]. The present research has stated that the consumption of user context in the event includes the history and features of user comments for improving the efficiency of CB classification or detection [11]. SSA with the DBN algorithm was used. Salp swarm algorithm (SSA) with deep belief network (DBN) is called as the SSA-DBN model. The SSA-DBN model is employed to detect and classify cyberbullying in social networks. For identifying suspicious attacks in a social, a salp swarm algorithm-based deep belief network is presented. As a result, the suggested chronological salp swarm algorithm-based deep belief network is constructed by fusing the chronological and salp swarm concepts. The fitness function, which accepts the minimal error vale as the optimal solution, reveals the optimal solution for detecting the incursion. The suggested approach tunes the weights appropriately in this case to produce an effective and optimal solution for identifying intruders.

This study presents a novel feature subset selection with DL-based CB detection and classification (FSSDL-CBDC) model on social networks. In addition, a binary coyote optimization-based feature subset selection (BCO-FSS) technique is applied to choose a set of features for enhanced classification efficiency. Moreover, the salp swarm algorithm (SSA) with deep belief network (DBN), called the SSA-DBN model, is working to detect and classify CB in social networks. Deep belief networks (DBNs) were created as a response to the issues that classic neural networks have with deep layered networks' training, such as slow learning, becoming stuck in local minima owing to poor parameter selection, and requiring much training datasets. The greedy algorithm is used to precondition deep belief networks—the design of the BCO-FSS and SSA-DBN models for CB detection and categorization procedure. To choose a set of features for improved classification efficiency, the binary coyote optimization-based feature subset selection (BCO-FSS) technique is used. To detect and classify CB in social networks, we combine the salp swarm algorithm (SSA) with

a deep belief network (DBN) and dubbed the SSA-DBN model. BCO-FSS and SSA-DBN model development for CB detection and classification process demonstrates the effort's inventiveness. Furthermore, the utilization of the SSA to fine-tune the hyperparameter of the DBN model resulted in enhanced outcomes over the traditional DBN model. The BCO technique is applied to choose a set of features for enhanced classification efficiency, and the SSA is employed to detect and classify cyberbullying in social networks. For the exploratory better detection presentation of the proposed FSSDL-CBDC method, a comprehensive range of simulations was performed on a benchmark dataset.

2. Related Works

This section reviews the recently developed automated CB classification models on social networks. Yuvaraj et al. [12] integrate the classification and feature extraction engine. The classification engine utilizing ANN categorizes the result, and it is given by a calculation scheme that either penalizes/rewards the categorized output. DRL performs the calculations, which increases the efficiency of classification. In their study, Mahbub et al. [13] investigate the impact of predatory approach words on CB detection and present a method for generating a vocabulary of predatory approach phrases. This study brings together findings from investigations of convicted criminals' chat logs in order to develop a lexicon of sexual approach terms for use in the future. Through the examination of data from a variety of social networks, the research establishes the relevance of this dictionary of approach terms in detecting online predatory behavior via machine learning methodologies. The variety of contents available on different social media sites are demonstrated by this example.

Talpur and O'Sullivan [14] created a supervised machine learning strategy for detecting CB and categorizing its severity on Twitter, which they published in Nature. The text classification engine created by Yuvaraj et al. [15] that preprocesses tweets, eliminates noisy data and other background information, extracts the desired features, and categorizes without overfitting the data is described in detail below. This research advances a novel DDT strategy that processes input components by utilizing the DNN hidden layer as tree nodes, as demonstrated in previous research. Chia et al. [16] use feature engineering and machine learning approaches to explore the use of irony and sarcasm on social media platforms. To begin, they define and assess the definitions of sarcasm and irony by looking at a large number of research studies that are focused on the contexts in which they are used. Subsequently, a comparison of numerous classification approaches with a few widely used classification schemes for the text classification process is carried out following the initial research. A variety of methods of data preprocessing were examined and compared in the following research.

In Murnion et al. [17], an automated data collection scheme is proposed that always gathers game chat data from the common online multiple player games. The data have been combined and collected using other data regarding the companies from the presented connected data service. It presented a scoring system for enabling the detection of CB depending upon this study. The organization of the gathered data was executed by humble feature recognition with SQL database enquiries and related with classifications from the AI-based sentimentality text analysis services, which have currently turned into presented and automatically classified data utilizing custom-built classification user.

Bu and Cho [18] proposed an ensemble technique of the 2 DL methods: first is character-level CNN that takes lower-level syntactic data from the series of characters and strong to the noise by the TL method. Next is word-level LRCN that takes higher-level semantic data from the series of words, accompanying the CNN module. Kumari and Singh [19] extract integrated features of text and images for identifying distinct events of CB. They utilize a pretrained VGG-16 network and CNN for extracting the features from text and images, correspondingly. These features are additionally improved by GA for increasing the performance of the entire system. Al-Garadi et al. conducted an in-depth research on cyberbullying prediction models on social media and identified several unresolved issues, including the prediction of cyberbullying intensity, human data features, and language dynamics. Numerous studies examine various machine learning options for detecting cyberbullying. Hosseinmardi et al. investigate the detection of cyberbullying episodes on the social media platform Instagram. They employ naive Bayes and SVM classifiers, the latter of which achieves the highest performance by combining multimodal text and image information, and media session data. Several other studies concentrate on the characteristics believed to be associated with cyberbullying, such as analyzing the social network structure of users, combining text and picture analysis techniques, profanity features, sentiment analysis, or geographical features, among others.

3. The Proposed Model

As illustrated in Figure 1, the operating principle of the FSSDL-CBDC approach is described. Several steps are required for the proposed FSSDL-CBDC technique to be effective, including preprocessing, feature selection using BCO-FSS, and classification using SSA-DBN. The FSSDL-CBDC technique was used. A variety of simulations were run to demonstrate the proposed FSSDL-CBDC technique's improved classification performance. Cyberbullying is a pernicious form of online abuse of authority that has malicious consequences. It takes on a variety of formats, and in the majority of social media platforms, it is in textual format. Intelligent systems are required to automatically detect such situations. Deep learning-based models have made their way into the detection of cyberbullying occurrences, claiming to be able to overcome the limits of conventional models and significantly enhance detection performance. The activities of these processes are described in greater detail in the following sections.
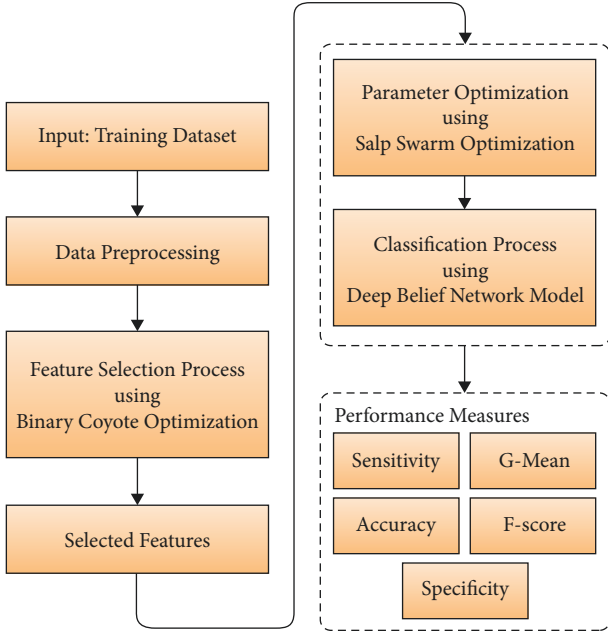
Figure 1: Working process of the FSSDL-CBDC model.

3.1. Preprocessing. During preprocessing, a lexical normalization technique is employed, which makes use of different elements for cleaning the input data. It also transforms the numerical parameters to the corresponding textual data. In addition, a spell corrector tool is used to eradicate the outbound vocabulary words. In addition, the repetitive or missing parameters are removed involving spelling mistakes, incorrect punctuation marks, and so on.

3.2. Design of the BCO-FSS Technique. Once the input social network data are preprocessed, they are fed into the BCO-FSS technique. The COA is a strong populace-based method projected lately by Juliano and Leandro [20]. This method draws stimulation from the common behaviors of *Canis latrans* species that live mostly in the NA. Because of its exclusive form, the COA could be categorized as evolutionary heuristics and swarm intelligence (SI). The coyote's population is separated to N_p packs, and N_c coyotes for each pack. The amount of coyotes in apiece pack is thought to be constant and equal. As a result, the populace size may be calculated by multiplying N_p and N_c. Every coyote's social status denotes a potential solution $\overrightarrow{x}$ to the optimization issue. In this case, the communal situation of cth coyote in p^{th} packet at t^{th} time may be described as follows:

$$soc_c^{p,t} = \overrightarrow{x} = (x_1, x_2, \ldots, x_D), \quad (1)$$

where D represents the amount of dimension in the search space. Initially, a population of $N_p \times N_c$ coyote is arbitrarily initiated within the predetermined search space as follows:

$$soc_{cj}^{p,t} = lb_j + r_j \cdot (ub_j - lb_j), \quad (2)$$

where lb and ub denote lower and upper bounds of jth decision variable, correspondingly. r_j represents real number arbitrarily created among zero and one, succeeding

uniform distributions. Then, the adaption of coyotes to their corresponding social condition is estimated by the following:

$$fit_c^{p,t} = f(soc_c^{p,t}). \quad (3)$$

As mentioned, the coyotes tend to leave their present pack for leading a lonely life or join other packs. The exclusion of coyote from its pack shadows a likelihood P_e, which differs based on the pack size as follows:

$$P_e = 0.005 \cdot N_c^2. \quad (4)$$

This method improves population range by stimulating a global interchange of data between the coyote packs. As P_e goes beyond one for $N_c \geq \sqrt{200}$, the COA limits the maximal amount of coyotes for each pack to fourteen [21]. In every pack, the coyote is optimally familiarized with the atmosphere and is allocated as alpha. For the minimization problem, the alpha is given as

$$alpha^{p,t} = \left\{ soc_c^{p,t} | arg_c = \{1, 2, \ldots, N_c\} \min f(soc_c^{p,t}) \right\}. \quad (5)$$

With the consideration of clear indications of the SI in coyotes, the COA assumes all the coyotes' share their social condition with the remaining packs for improving the pack's survivability. Regarding this, the traditional tendency of a pack is determined according to the data given its member, ciz.,

$$cult_j^{p,t} = \begin{cases} O_{(Nc+1)/2}^{p,t} & N_c \text{ is odd} \\ \dfrac{O_{(Nc/2),j}^{p,t} + O_{(Nc+1)/2,j}^{p,t}}{2} & \text{otherwise} \end{cases}, \quad (6)$$

where $O^{p,t}$ represents hierarchical communal condition of the prairie wolf within p^{th} pack at t^{th} time, for $j = 1, 2, \ldots, D$. In this method, the traditional propensity of every pack is calculated as the middle of the combined communal condition within the carton. For modeling the two main organic proceedings of the coyote, namely, birth and death, the ages of every coyote are deliberated, age $\in$ N. The birth of novel coyotes is defined by the combination among the communal circumstances of the two arbitrary parental coyotes from similar packs, and the effect of the environmental factor R_j,

$$pup_j^{p,t} = \begin{cases} soc_{r1,j}^{p,t}, & r \text{ and } j < P_a \text{ or } j = j_1 \\ soc_{r2,j}^{p,t}, & r \text{ and } < P_s + P_a \text{ or } j = j_2, \\ R_j, & \text{otherwise} \end{cases} \quad (7)$$

where $soc_{r1,j}^{p,t}$ and $soc_{r2,j}^{p,t}$ denote arbitrary coyote from the pth pack, and j_1 and j_2 denote two arbitrary dimensions of the search space. Alternatively, P_s and P_a denote scatter and relationship likelihoods, correspondingly. R_j denotes arbitrarily created vector within the bounds of jth dimension, and r and $_j$ denotes uniformly arbitrary amounts in zero and one. The scatter and relationship likelihoods have a substantial impact on the composition and diversity of the coyote's pack. It is given by the following:

$$P_s = \frac{1}{D},$$

$$P_a = \frac{(1 - P_s)}{2}. \tag{8}$$

Based on this, the coyote's pup has around 10% chance of death at birth. Moreover, the death risk of every coyote increases with age. Thus, the COA designs the coyote's survivability depending upon a simple method, whereas ω and φ denote set of coyotes fewer adapted to the atmosphere (viz. worst fitness value) compared to pup and size of the groups, correspondingly.

Additionally, the COA represents the traditional communication among the coyotes in the pack using δ_1 and δ_2. The previous denotes the effect of alpha on an arbitrary coyote cr_1, where the later indicates the effect of the traditional tendency of the pack on other arbitrary coyotes cr_2. The cr_1 and cr_2 are chosen after a uniformly distributed likelihood. Therefore, δ_1 and δ_2 are given by

$$\delta_1 = \text{alpha}^{p,t} - \text{soc}_{cr1}^{p,t}, \tag{9}$$

$$\delta_2 = \text{cult}^{p,t} - \text{soc}_{cr2}^{p,t}. \tag{10}$$

The social condition of the coyote is inclined by the alpha, and other members of the pack will be upgraded using

$$\text{new soc}_c^{p,t} = \text{soc}_c^{p,t} + r_1 \cdot \delta_1 + r_2 \cdot \delta_2, \tag{11}$$

where r_1 and r_2 are random values. The coyotes of new social conditions are assessed using

$$\text{new fit}_c^{p,t} = f\left(\text{new} - \text{soc}_c^{p,t}\right). \tag{12}$$

The coyotes' general fitness will either increase or remain the same, but it will never deteriorate.

$$\text{soc}_c^{p,t+1} = \begin{cases} \text{new\_soc}_c^{p,t} & \text{if new} - \text{fit}_c^{p,t} < \text{fit}_c^{p,t} \\ \text{soc}_c^{p,t} & \text{otherwise} \end{cases}. \tag{13}$$

Additional features could inhibit the learning procedure. The FS could authenticate the significance of the features, which create a dataset, and removing these does not assist in a positive manner. The selected features using an FS method could be denoted as N-sized vector, whereas N denotes the overall number of features in a dataset, whereas every position of the vector could consider the values as zero/one, whereas zero denotes features that were not chosen as also one signifies features that were chosen. The transfer purpose technique determines the likelihood of altering a position vector component from zero to one and conversely in an efficient and simple manner, and hence, the binarization method is the utilized most, particularly for the FS problem [22]. Based on this, a transfer function considerably affects the efficiency of the FS methods in seeking an optimal set of features, concerning the local optimal prevention and the balance between exploitation and exploration, and thus, it is a significant role in the binary version of metaheuristics. In the BCOA, the constraint of the social conditions of coyotes

for the binary values by a V-shaped transfer function is defined by equation (14) as follows:

$$V\left(\text{new\_soc}_c^{p,t}\right) = \left| \frac{\text{new\_soc}_c^{p,t}}{\sqrt{1 + \left(\text{new\_soc}_c^{p,t}\right)^2}} \right|, \tag{14}$$

where $\text{new\_soc}_c^{p,t}$ relates to the upgraded social condition vector existing, considering continuous values.

3.3. Algorithmic Design of the SSA-DBN Technique. Once the subsets have been feature-reduced, they are fed into the SSA-DBN model, which is then utilized to complete the classification assignment. This is accomplished through the use of the DBN model, which generates feature vectors that are then classified using the softmax layer. By picking hyperparameter values in the most optimal method, the SSA is used to improve the detection performance of a DBN model, whereas the SSA is used to improve its detection performance.

3.3.1. Architecture of DBN. Since DBN contains multiple hidden levels and countless hidden units within each of those layers, it is considered a member of the DNN family. The standard DBN technique is similar to the RBM technique in that it includes an output layer. Additionally, the DBN achieves its outcomes using a strong, greedy unsupervised learning technique to train RMB, and a supervised fine-tuning mechanism that adjusts the scheme using labeled data. The RBM is composed of two types of layers: visible and buried layers coupled with undirected weights. When RBMs are stacked in DBN, the RBM's hidden layer is chosen to be the visible layer of the subsequent RBM. This is because the RBM's hidden layer provides information about the subsequent RBM. RMB's variable sets are defined as $= (w, b, a)$, where w ij is the weight difference between v I and h j. The bias layers b I and a j must be identified.

Figure 2 displays the framework of DBN. The RBM defines equivalent energy as described as follows:

$$E(v, h \mid \theta) = -\sum_i b_i v_i - \sum_j a_j h_j - \sum_i \sum_j w_{ij} v_i h_j. \tag{15}$$

The joint likelihood distribution of v and h is defined below as follows:

$$p(v, h \mid \theta) = \frac{\exp(-E(v, h \mid \theta))}{\sum_{v,h} \exp(-E(v, h \mid \theta))}. \tag{16}$$

Here, the marginal likelihood distribution of v is established by

$$p(v \mid \theta) = \frac{\sum_h \exp(-E(v, h \mid \theta))}{\sum_{v,h} \exp(-E(v, h \mid \theta))}. \tag{17}$$

To gain an optimum θ value for a solitary data vector v, the incline of log probability approximation is evaluated by [23] the following:
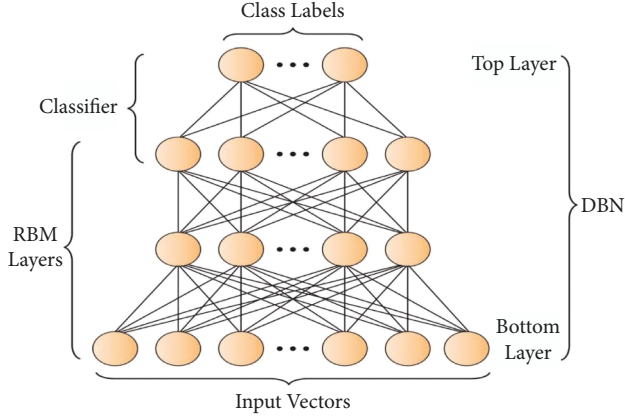
Figure 2: DBN structure.

$$\frac{\partial \log\ p(v\,|\,\theta)}{\partial w_{ij}} = v_i h_{j\,\text{data}} - v_i h_{j\,\text{model}},$$

$$\frac{\partial \log\ p(v\,|\,\theta)}{\partial a_j} = h_{j\,\text{data}} - h_{j\,\text{model}}, \qquad (18)$$

$$\frac{\partial \log\ p(v\,|\,\theta)}{\partial b_i} = v_{i\,\text{data}} - v_{i\,\text{model}}.$$

Here, $\langle\bullet\rangle$ denotes expectation using the delivery of a specific subscript. Because of the lack of influences between units within the same layer, $\langle\bullet\rangle$ _data suggests that it may be obtained by measuring the conditional likelihood distribution by measuring the provisional likelihood delivery by

$$p\left(h_j\,|\,v,\,\theta\right) = \frac{1}{1 + \exp\left(-\sum_j w_{ij} v_i - a_j\right)},$$

$$p\left(v_i\,|\,h,\,\theta\right) = \frac{1}{1 + \exp\left(-\sum_j w_{ij} h_j - b_i\right)}. \qquad (19)$$

Due to the shape of the activation function, it is referred to as a sigmoid function. Contrastive divergence can be regarded of as a learning technique that approximates maximum likelihood. It calculates the divergence/differences between the positive phase (energy of first encoding) and the negative phase (energy of second encoding) (energy of the last encoding). To reduce the variation of two Kullback-Leibler divergences through renovation, the contrastive divergence (CD) learning module is used in the case of _model (KL). To begin, the CD learning is more efficient than Gibbs sampling in real-world applications and requires less processing time. Thus, weights in the DBN layer are taught using unlabeled data by unsupervised algorithms that are both fast and greedy in their information search. For predictions, the DBN uses the supervised layer to fine-tune the learned features using labeled data from the training set. The fully connected (FC) layer

is now the top layer, and the layers beneath it are activated using the sigmoid activation function.

3.3.2. Overview of SSA. The SSA is a novel SI method, which was established lately by [24]. The main concept behindhand in the SSA operator is that they simulate the swarming behavior of salp in deep oceans. Salp belonged to the species of Salpidae and contain transparent barrel-shaped bodies. They are related to jellyfishes in their tissue and motion. Moreover, they shift as the water is driven by the body as a force to move onward. The salp provides a novel 160 forms of swarm called a slapping chain while directing in the ocean. The salp chain behavior was numerically modeled by separating the population into groups depending upon leader and follower. The front of the chain is deliberated as the led 1 when the remaining salps are called as followers. The leader's part is to direct the swarm of salp, and all the followers follow the previous one. Related to the other SI techniques, the procedure of SSA imitates by initiating an arbitrary population of salp and later evaluates the fitness for every salp [25]. The slap with an optimal best fitness value is represented as a front-runner salp, whereas an additional slap is symbolized as a follower. The salp swarm algorithm (SSA) is a new stochastic algorithm inspired by salps' navigational and foraging abilities. However, higher dimension problems show a poor convergence rate for classical SSA. The SSA lacks exploration and exploitation, resulting in inefficient convergence. A salp's best fitness is found by exploring and exploiting the search space. The leader's salp location is modified based on the distance between the salp and food supply. The optimal performing slaps are represented as a food source to be chased using a salp chain. For updating the location of the slap chain, two major stages are determined: leader and follower phases.

The location of the leader is upgraded by equation (20) as follows:

$$X_j^1 = \begin{cases} X_{\text{Best}_j} + c_1\left(\left(ub_j - lb_j\right)c_2 + lb_j\right) \text{if } c_3 \geq 0.5, \\ X_{\text{Best}_j} - c_1\left(\left(ub_j - lb_j\right)c_2 + lb_j\right)\text{else}, \end{cases} \qquad (20)$$

where X_j^1 and X_{Best_j} represent novel location of leader and food source in *jth* dimension, and ub_j and lb_j represent upper and lower bounds of jth dimension, correspondingly. c_2 and c_3 denote arbitrarily created amounts in the range zero and one . The variable c_1 presents an important aspect in the SSA that controls the balance between exploration and exploitation. Moreover, c_1 gradually decreases by iteration as displayed in the following equation (21):

$$c_1 = 2e^{-(4t/T)^2}, \qquad (21)$$

where t specifies present iteration, and T denotes maximal amount of 152 iterations. Figure 3 exemplifies the flowchart of the SSA. For updating the location of follower, novel idea
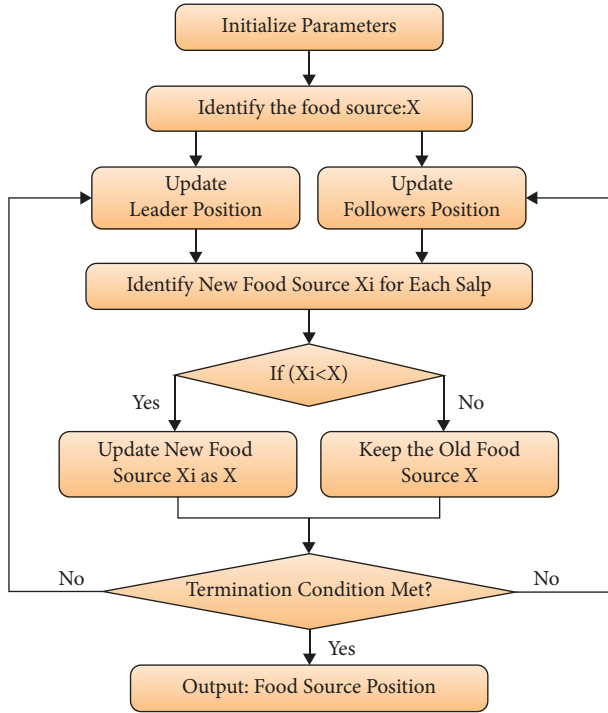
Figure 3: Flowchart of SSA.

is presented where 185 is depending upon Newton's law of movement as in equation (22).

$$X_j^i = \frac{1}{2} g t^2 + w_0 t, \ i \geq 2, \tag{22}$$

where X_j^i denotes location of ith follower salp in j^{th} dimension. In the optimization procedure, the time t corresponds to the present iteration, 188, whereas g and w_0 indicate velocity and acceleration, correspondingly. In equation (22), the early speed w_0 is set to zero and the inconsistency is set 190 to one ($\triangle t = 1$); thus, the upgrading procedure of follower is equated in 191 equation (23).

$$X_j^i j = \frac{1}{2} \left(X_j^i + X_j^{i-1} \right). \tag{23}$$

3.3.3. Parameter Optimization of DBN Using SSA. For optimally regulating the hyperparameters of the DBN model, the SSA is used and the detailed working is provided in the following. The training process of the DBN model takes place using a fitness function [26–30]. In addition, 10-fold cross-validation (CV) process is utilized to evaluate the FF. Under 10-fold CV, the training dataset is randomly subdivided into a collection of ten equally exclusive subsets of nearly equal sizes, where nine subsets are used to train the data, and the remaining one is applied to test the data. These processes are repeated for a set of 10 iterations in such a way that each subset can be used to test the model. The FF is denoted as 1-$CA_{\text{validation}}$ of the 10-fold CV model in the training data, as defined in equation (24). Also, a solution with maximum $CA_{\text{validation}}$ leads to minimal fitness value [31–33].

$$\text{Fitness} = 1 - CA_{\text{validation}},$$

$$CA_{\text{validation}} = 1 - \frac{1}{10} \sum_{i=1}^{10} \left| \frac{y_c}{y_c + y_f} \right| \times 100, \tag{24}$$

where y_c and y_f indicate the true and false organization count. Finally, the hyperparameter involved in the DBN model is optimally picked up by the SSA, and also, the performance of classification gets improved.

4. Performance Evaluation

In this section, we validate the proposed model performance under several aspects. Table 1 investigates the performance of the feature selection techniques in terms of classification accuracy under different sets of training data and varying number of residuals [34, 35]. Figure 4 examines the result analysis of different feature selection techniques in terms of classification accuracy on 60% of training data. From the figure, it is depicted that the BCO-FSS model is found to be an effective method and it leads to maximum classification accuracy. For instance, under 200 residuals, the proposed BCO-FSS technique has accomplished a higher classification accuracy of 28.61%, whereas Pearson's correlation, chi-squared, and information gain techniques have resulted in a lower classification accuracy of 28.61%, 27.10%, and 25.18%, respectively.

Moreover, under 1000 residuals, the BCO-FSS technique has obtained an increased classification accuracy of 37.80%, whereas Pearson's correlation, chi-squared, and information gain techniques have attained a decreased classification accuracy of 34.78%, 32.68%, and 29.46%, respectively.

Figure 5 inspects the outcome analysis of different feature selection approaches with respect to classification accuracy on 75% of training data. From the figure, it can show that the BCO-FSS method is found to be an effective technique and it leads to maximal classification accuracy. For instance, under 200 residuals, the presented BCO-FSS manner has accomplished a higher classification accuracy of 43.24%, whereas Pearson's correlation, chi-squared, and information gain methods have resulted in a lesser classification accuracy of 40.39%, 36.26%, and 34.65%, correspondingly. Furthermore, under 1000 residuals, the BCO-FSS technique has gained an improved classification accuracy of 66.26%, whereas Pearson's correlation, chi-squared, and information gain techniques have achieved a reduced classification accuracy of 60.07%, 55.47%, and 52.19%, correspondingly.

Figure 6 investigates the result analysis of different feature selection methods in terms of classification accuracy on 90% of training data. From the figure, it can be stated that the BCO-FSS manner is initiated to be an effective approach and it leads to higher classification accuracy. For sample, under 200 residuals, the projected BCO-FSS method has talented a larger classification accuracy of 47.44%, while Pearson's correlation, chi-squared, and information gain techniques have resulted in a lower classification accuracy 43.88%, 39.55%, and

TABLE 1: Consequences of existing with future feature assortment methods on various training data size.

Classification accuracy on 60% training data			
Residuals	BCO-FSS	Pearson correlation	Chi-squared	Information gain
200	28.61	27.10	25.18	23.21
400	31.87	30.47	28.26	26.54
600	33.09	31.23	29.43	27.41
800	35.30	33.15	31.29	28.97
1000	37.80	34.78	32.68	29.46
Classification accuracy on 75% training data			
Residuals	BCO-FSS	Pearson correlation	Chi-squared	Information gain
200	43.24	40.39	36.26	34.65
400	53.24	50.07	44.36	41.84
600	55.94	51.66	47.85	43.75
800	61.02	56.10	51.66	48.06
1000	66.26	60.07	55.47	52.19
Classification accuracy on 90% training data			
Residuals	BCO-FSS	Pearson correlation	Chi-squared	Information gain
200	47.44	43.88	39.55	33.29
400	61.45	56.10	49.23	44.62
600	66.80	57.88	45.92	40.90
800	67.82	60.69	52.28	48.47
1000	74.18	69.60	64.51	59.61
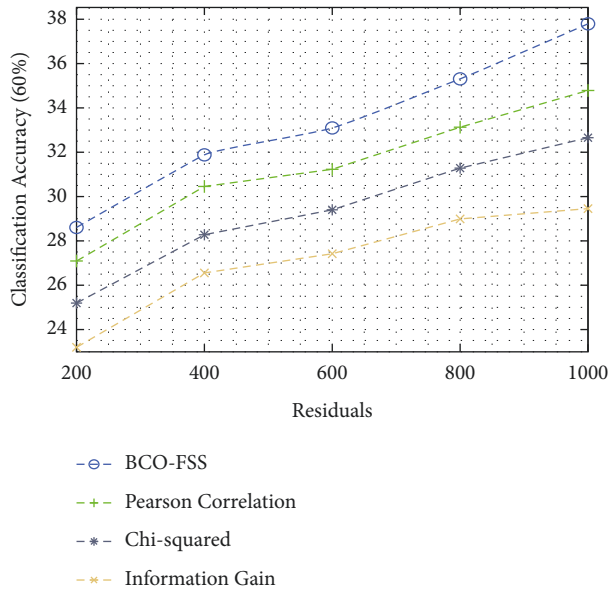
FIGURE 4: Classification accuracy analysis of the BCO-FSS model on 60% of training data.

FIGURE 5: Classification accuracy analysis of the BCO-FSS model on 75% of training data.
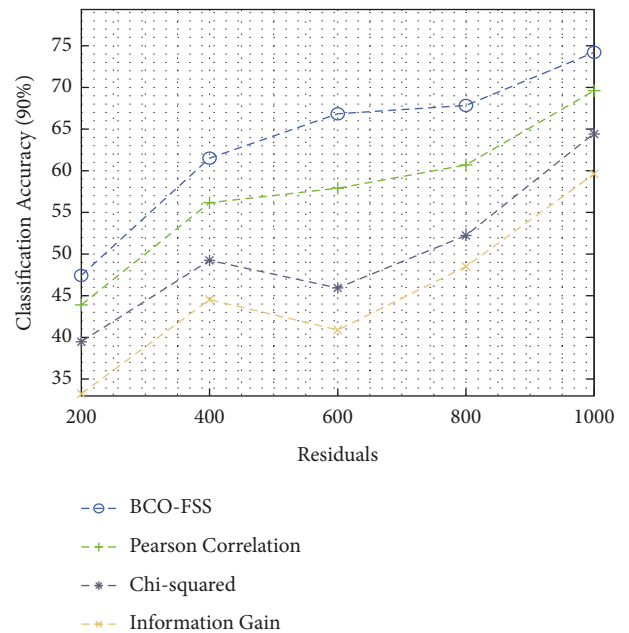
FIGURE 6: Classification accuracy analysis of the BCO-FSS model on 90% of training data.

33.29%, correspondingly. Also, under 1000 residuals, the BCO-FSS technique has attained a maximum classification accuracy of 74.18%, whereas Pearson's correlation, chi-squared, and information gain techniques have gained a lesser classification accuracy of 69.60%, 64.51%, and 59.61%, correspondingly.

If fewer than 60% of the training data are available, the suggested SSA-DBN model is compared to existing techniques in Table 2. Figure 7 displays an overview of the SSA-DBN model's sensitivity and specificity studies, which were conducted using the comparable methodologies. The LR model performed worse than the LR model, as seen in the figure, with a sensitivity of 74.335 % and a specificity of 75.025 %, respectively. It was expected that the RF model would produce a little better result, with a sensitivity of

TABLE 2: Results of existing with the proposed SSA-DBN on predicting cyberbullying with 60% training data.

Measures	NB	LR	RF	SVM	ANN	ANN-DRL	SSA-DBN
Accuracy	61.821	69.023	71.814	77.236	80.897	85.369	88.473
F-measure	72.674	72.965	73.045	73.325	77.876	83.588	86.901
G-mean	73.115	73.065	74.895	76.856	79.427	82.008	84.371
Sensitivity	80.657	74.335	74.825	76.506	76.976	83.598	85.728
Specificity	73.475	75.025	78.196	83.708	84.719	85.129	88.728
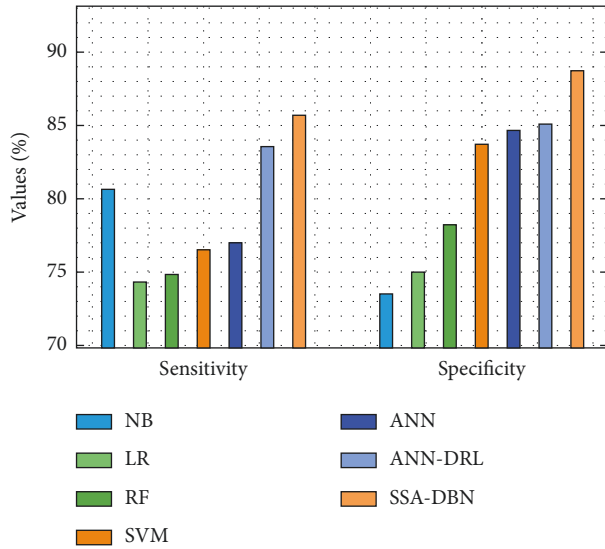
FIGURE 7: Sensitivity and specificity analysis of the SSA-DBN model on 60% training data.
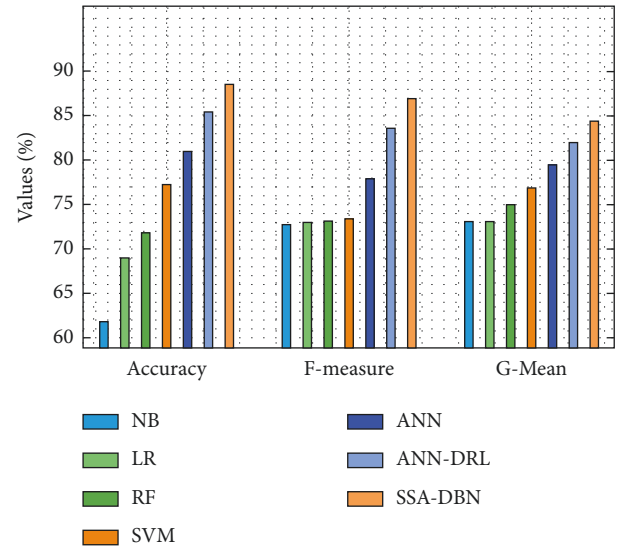
FIGURE 8: Comparative analysis of the SSA-DBN model on 60% training data.

74.825 % and a specificity of 78.196 %, and the RF model did so. Furthermore, the SVM model displayed significantly improved performance, with a sensitivity of 76.506 % and a specificity of 83.708 %, respectively, compared to the baseline model. The ANN model also generated moderate results, with a sensitivity of 76.976 % and a specificity of 84.719 %, respectively, for the sensitivity and specificity tests. Aside from that, the NB model provided a tolerable outcome, as evidenced by its sensitivity of 80.657 % and specificity of 73.475 %, among other statistics. Additionally, the ANN-DRL model must yield competitive outcomes with 83.598 percent% sensitivity and 85.129 % specificity in order to be considered successful. But the proposed SSA-DBN model outperformed the competition, with sensitivity and specificity values of 85.728 % and 88.728 %, respectively, in the study.

Figure 8 offers a comparative analysis of the SSA-DBN with other classifiers in terms of accuracy, F-measure, and G-mean. The obtained results illustrated that the NB and LR models have attained a lower accuracy of 61.821 and 69.023, respectively. Next, the RF and SVM models have showcased moderately closer presentation with the correctness of 71.814% and 77.236%, respectively. In addition, the ANN and ANN-DRL techniques have showcased reasonable consequences with the correctness of 80.897% and 85.369%, correspondingly. However, the SSA-DBN perfect has outperformed the other methods with a maximum accuracy of 88.473%.
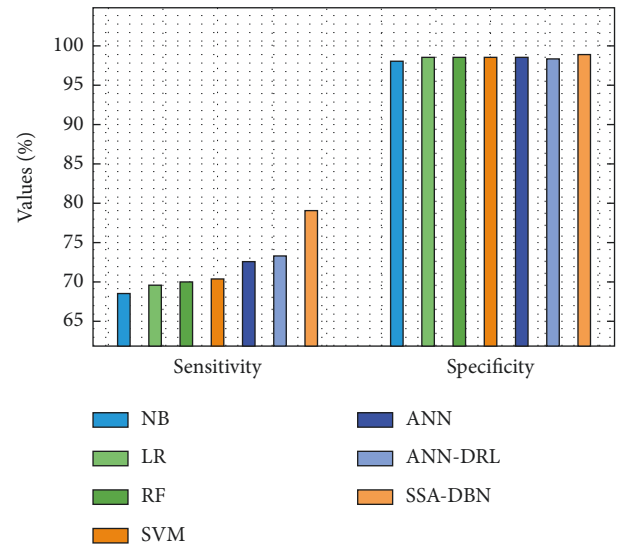
FIGURE 9: Sensitivity and specificity analysis of the SSA-DBN model on 75% training data.

On the following page, you will find a quick comparison of the given SSA-DBN approach to alternative strategies that use less than 75 % of the training data. Figure 9 displays a detailed sensitivity and specificity evaluation of the SSA-

TABLE 3: Results of existing with proposed SSA-DNN on predicting the cyberbullying with 75% training data.

Measures	NB	LR	RF	SVM	ANN	ANN-DRL	SSA-DBN
Accuracy	96.743	97.063	96.993	97.003	97.553	97.313	98.362
F-measure	61.281	62.901	63.491	63.891	65.242	65.552	70.466
G-mean	82.238	82.498	83.388	83.468	84.499	85.189	88.471
Sensitivity	68.593	69.583	70.074	70.394	72.614	73.315	79.063
Specificity	98.093	98.583	98.513	98.663	98.553	98.353	98.976
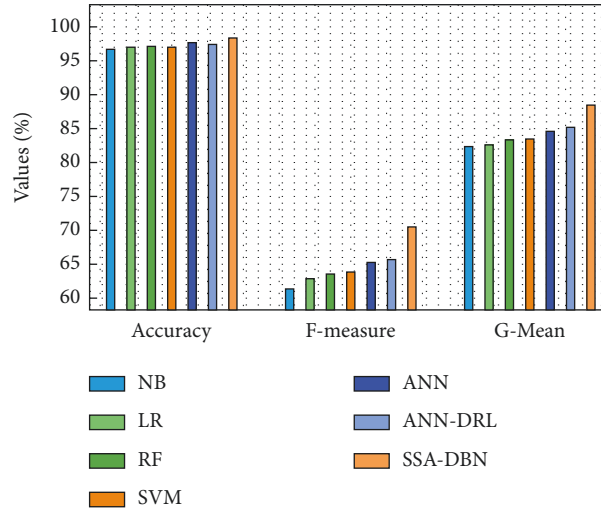
FIGURE 10: Comparative analysis of the SSA-DBN model on 75% training data.

DBN model, which was carried out utilizing comparative methodology. In this study, we discovered that the NB approach had lower sensitivity than the other methods (68.593 %) but greater specificity than the other methods (98.093 %). Additionally, the LR strategy produced somewhat better outcomes, with a sensitivity of 69.583 % and a specificity of 98.583 %, respectively, compared to the LR approach. Afterward, the RF algorithm displayed significantly improved performance, with a sensitivity of 70.074 % and a specificity of 98.513 %, respectively, compared to the baseline algorithm. The SVM model also performed well in terms of sensitivity and specificity, scoring 70.394 % and 98.663 %, respectively, for mild outcomes in the study. As an additional benefit of using the artificial neural network model, a tolerable outcome was achieved with a sensitivity of 72.614 % and a specificity of 98.553 %. At the same time, the ANN-DRL model aimed to exhibit competitive outcomes with 73.315 % compassion and 98.353 % specificity while also attempting to demonstrate competitive outcomes. However, the accessible SSA-DBN practice obtained optimal performance, with a compassion rate of 79.063 % and a specificity rate of 98.976 %, respectively, compared to the other practices tested in Table 3.

Figure 10 provides a comparative analysis of the SSA-DBN with other classifiers with respect to accuracy, F-measure, and G-mean. The attained outcomes showcased that the NB and RF methods have attained a lower accuracy of 96.743 and 96.993, correspondingly. Next, the SVM and LR techniques have depicted moderately closer performance

with the accuracy of 97.003% and 97.063%, correspondingly. Also, the ANN and ANN-DRL techniques have exhibited reasonable consequences with the correctness of 97.553% and 97.313%, respectively. However, the SSA-DBN perfect has outperformed the other techniques with a maximal accuracy of 98.362%.

With less than 90 % of the training data, the anticipated SSA-DBN model's performance is compared to the performance of other techniques, which is presented in detail in Table 4. According to the comparison methodologies used in this work, the SSA-DBN model's sensitivity and specificity were briefly evaluated (Figure 11). In the figure, it can be seen that the LR technique fared poorer than the other approaches, with a sensitivity of 92.766 % and a specificity of 99.583 %. Although the NB technique produced a little better results than expected (sensitivity 92.796 % and specificity 99.673 %), it did so in a more consistent manner. A similar improvement in performance was seen in the RF model, which had a sensitivity of 93.537 % and a specificity of 99.773 % when compared to the baseline model. The SVM model also showed moderate results, with a sensitivity of 93.697 % and a specificity of 99.803 %, respectively, according to the results of the study. In addition, the ANN model provided a tolerable outcome, with a sensitivity of 94.387 % and a specificity of 99.733 %, according to the results. As part of this effort, the ANN-DRL approach aimed to simultaneously exhibit competitive outcomes. The method's sensitivity and specificity were both 99.873 %, indicating that it was successful in demonstrating competitive outcomes. The new

TABLE 4: Results of existing with the proposed SSA-DNN on predicting cyberbullying with 90% of training data.

Measures	NB	LR	RF	SVM	ANN	ANN-DRL	SSA-DBN
Accuracy	99.503	99.633	99.623	99.673	99.563	99.703	99.983
F-measure	88.071	88.191	90.213	90.063	91.454	91.455	94.378
G-mean	96.299	96.269	96.530	96.640	97.030	97.000	98.361
Sensitivity	92.796	92.766	93.537	93.697	94.387	94.418	96.037
Specificity	99.673	99.583	99.773	99.803	99.733	99.873	99.997
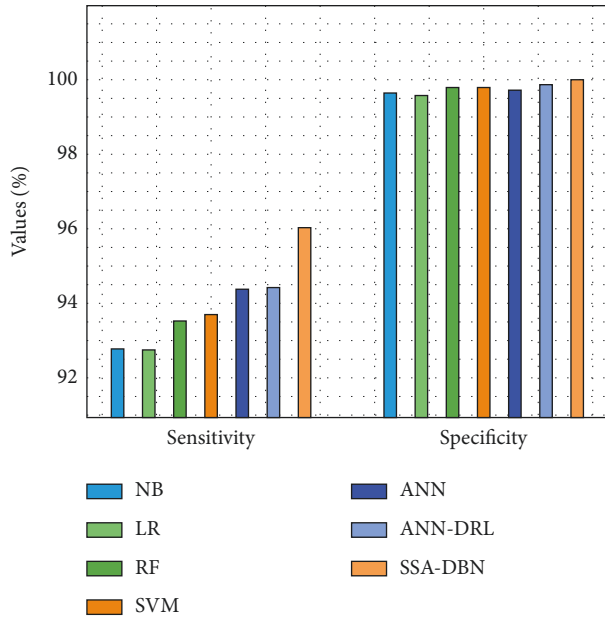
FIGURE 11: Compassion and specificity analysis of the SSA-DBN model on 90% of training data.
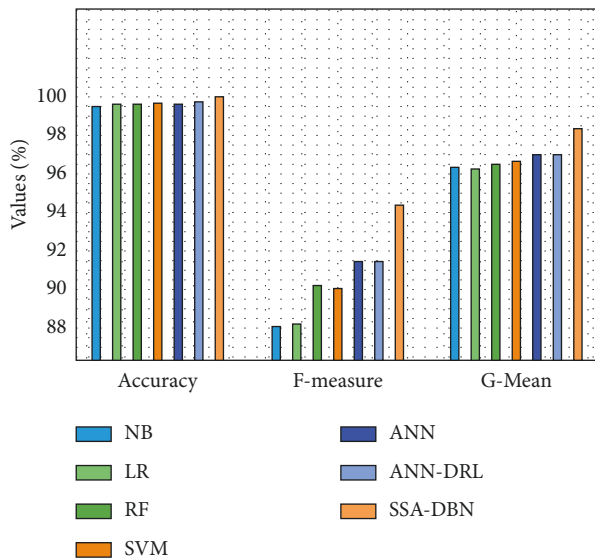
FIGURE 12: Comparative analysis of the SSA-DBN model on 90% training data.

SSA-DBN model, on the other hand, outperformed the prior model, with sensitivity and specificity values of 96.037 and 99.997 %, respectively, in comparison.

Figure 12 gives a comparative analysis of the SSA-DBN with other classifiers in terms of accuracy, F-measure, and G-mean. The attained results demonstrated that the NB and RF approaches have attained a lower accuracy of 99.503 and 99.623, respectively. In addition, the LR and SVM copies have showcased abstemiously closer presentation with the correctness of 99.633% and 99.673%, correspondingly. At the same time, the ANN and ANN-DRL manners have showcased reasonable consequences with the correctness of 99.563% and 99.703%, congruently. However, the SSA-DBN perfect has demonstrated the other algorithms with a higher accuracy of 99.983%.

5. Conclusions

In this post, we will describe a novel FSSDL-CBDC technique for detecting and classifying cyberbullying in social media, and how to apply it. The suggested FSSDL-CBDC technique consists of a number of phases, including preprocessing, feature selection, and classification, among others. Additionally, by creating the BCO-FSS approach to choose the optimal collection of features from the preprocessed data, the overall classification results are significantly enhanced. Figure 1 shows the BCO-FSS technique design. The SSA-DBN model receives and classifies the feature-reduced subset in the same time frame as the other models. When compared to the classic DBN model, the usage of the SSA to fine-tune the DBN model's hyperparameter resulted in improved outcomes when using the SSA. A large number of simulations on a benchmark dataset were carried out in order to assess the increased detection performance of the proposed FSSDL-CBDC technique, which was found to be effective. When compared to other state-of-the-art approaches, the simulation results revealed that the FSSDL-CBDC strategy performed significantly better in classification than the others. In the future, the performance of the FSSDL-CBDC technique may be enhanced by including outlier identification and feature reduction techniques in the algorithm. Unsupervised feature selection (FS) for outlier detection (OD) in streaming data (SD) for fields such as intrusion detection and network security, which are increasingly challenged with large amounts of high-dimensional data that must be analyzed in near real time.

Data Availability

This article contains all of the data.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] S. R. Sahoo and B. B. Gupta, "Classification of various attacks and their defence mechanism in online social networks: a survey," *Enterprise Information Systems*, vol. 13, no. 6, pp. 832–864, 2019.

[2] R. Annamalai, S. J. Rayen, and J. Arunajsmine, "Social media networks owing to disruptions for effective learning," *Procedia Computer Science*, vol. 172, pp. 145–151, 2020.

[3] M. Fire, G. Katz, and Y. Elovici, "Strangers intrusion detection detecting spammers and fake profiles in social networks based on topology anomalies," *Human Journal*, vol. 1, no. 1, pp. 26–39, 2012.

[4] D. Paulraj, "A gradient boosted decision tree based sentiment classification of twitter data," *International Journal of Wavelets Multiresolution and Information Processing*, vol. 18, no. 4, pp. 2050027–2050121, 2020.

[5] H. Kefi and C. Perez, "Dark Side of Online Social Networks: Technical, Managerial, and Behavioral Perspectives," *Encyclopedia of Social Network Analysis and Mining*, vol. 143, pp. 535–556, 2018.

[6] D. Vinotha and M. Vasanthi, "Classification rule discovery using ant-miner algorithm: an application of network intrusion detection," *Nternational Journal Of Modern Engineering Research (IJMER)*, vol. 4, no. 8, 2014.

[7] S. Hinduja and J. W. Patchin, "Bullying, cyberbullying, and suicide," *Archives of Suicide Research*, vol. 14, no. 3, pp. 206–221, 2010.

[8] S. Tripathi, V. B. Devi, I. Bhardwaj, and N. Arulkumar, "IoT-based traffic prediction and traffic signal control system for smart city," *Soft Computing*, vol. 25, no. 18, pp. 12241–12248, 2021.

[9] M. W. Savage and R. S. Tokunaga, "Moving toward a theory: testing an integrated model of cyberbullying perpetration, aggression, social skills, and Internet self-efficacy," *Computers in Human Behavior*, vol. 71, pp. 353–361, 2017.

[10] P. Subbulakshmi and V. Ramalakshmi, "Honest Auction Based Spectrum Assignment," *Wireless Personal Communications*, vol. 102, no. 1, 2008.

[11] T. Vaillancourt, R. Faris, and F. Mishna, "Cyberbullying in children and youth: implications for health and clinical practice," *Canadian Journal of Psychiatry*, vol. 62, no. 6, pp. 368–373, 2017.

[12] N. Yuvaraj, K. Srihari, G. Dhiman et al., "Nature-inspired-based approach for automated cyberbullying classification on multimedia social networking," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6644652, 12 pages, 2021.

[13] S. Mahbub, E. Pardede, and A. S. M. Kayes, "Detection of Harassment Type of Cyberbullying: A Dictionary of Approach Words and its Impact," *Security and Communication Networks*, vol. 2021, Article ID 5594175, 12 pages, 2021.

[14] B. A. Talpur and D. O'Sullivan, "Cyberbullying severity detection: a machine learning approach," *PLoS One*, vol. 15, no. 10, Article ID e0240924, 2020.

[15] N. Yuvaraj, V. Chang, B. Gobinathan et al., "Automatic detection of cyberbullying using multi-feature based artificial intelligence with deep decision tree classification," *Computers & Electrical Engineering*, vol. 92, Article ID 107186, 2021.

[16] Z. L. Chia, M. Ptaszynski, F. Masui, G. Leliwa, and M. Wroczynski, "Machine Learning and feature engineering-based study into sarcasm and irony classification with application to cyberbullying detection," *Information Processing & Management*, vol. 58, no. 4, Article ID 102600, 2021.

[17] S. Murnion, W. J. Buchanan, A. Smales, and G. Russell, "Machine learning and semantic analysis of in-game chat for cyberbullying," *Computers & Security*, vol. 76, pp. 197–213, 2018.

[18] S. J. Bu and S. B. Cho, "A hybrid deep learning system of CNN and LRCN to detect cyberbullying from SNS comments," *Hybrid Artificial Intelligent Systems. HAIS 2018. Lecture Notes in Computer Science*, Springer, vol. 10870, pp. 561–572, 2018.

[19] K. Kumari and J. P. Singh, "Identification of cyberbullying on multi modal social media posts using genetic algorithm," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 2, Article ID e3907, 2021.

[20] J. P. L. d. S. Coelho, "Coyote Optimization Algorithm: a new metaheuristic for global optimization," in *IEEE Congress on Evolutionary Computation (CEC)*pp. 2633–2640, Rio de Janeiro, Brazil, 2018.

[21] S. Saravanan, M. Hailu, G. M. Gouse, M. Lavanya, and R. Vijaysai, "Optimized secure scan flip flop to thwart side channel attack in crypto-chip," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 274, pp. 410–417, 2019.

[22] R. C. T. de Souza, C. A. de Macedo, L. dos Santos Coelho, J. Pierezan, and V. C. Mariani, "Binary coyote optimization algorithm for feature selection," *Pattern Recognition*, vol. 107, Article ID 107470, 2020.

[23] J. Yu and G. Liu, "RETRACTED: knowledge-based deep belief network for machining roughness prediction and knowledge discovery," *Computers in Industry*, vol. 121, Article ID 103262, 2020.

[24] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm Algorithm a bio inspired optimizer for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163–191, 2017.

[25] S. Neelakandan and D. Paulraj, "An automated exploring and learning model for data prediction using balanced CA svm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 5, p. 4990, 2020.

[26] L. Abualigah, M. Shehab, M. Alshinwan, and H. Alabool, "Salp Swarm Algorithm: A Comprehensive Survey," *Neural Computing and Applications*, vol. 32, pp. 11195–11215, 2019.

[27] V. J. Chin and Z. Salam, "Coyote optimization algorithm for the parameter extraction of photovoltaic cells," *Solar Energy*, vol. 194, pp. 656–670, 2019.

[28] R. F. Mansour, N. M. Alfar, S. Abdel-Khalek, M. Abdelhaq, R. A. Saeed, and R. Alsaqour, "Optimal deep learning based fusion model for biomedical image classification," *Expert Systems*, vol. 39, no. 3, Article ID e12764, 2021.

[29] J. R. Beulah, L. Prathiba, G. L. N Murthy, E. Fantin Irudaya Raj, and N. Arulkumar, "Blockchain with Deep Learning Enabled Secure Healthcare Data Transmission and Diagnostic Model," *International Journal of Modeling, Simulation, and Scientific Computing*, Article ID 2241006, 2022.

[30] T. Kavitha, C. Karthikeyan, M. Ashok, R. Kohar, J. Avanija, and S. Neelakandan, "Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 1, pp. 113–129, 2021.

[31] D. Venu, A. V. R. Mayuri, S. Neelakandan, G. Murthy, N. Arulkumar, and N. Shelke, "An efficient low complexity compression based optimal homomorphic encryption for secure fiber optic communication," *Optik*, vol. 252, Article ID 168545, 2022.

[32] S. Neelakandan, M. Prakash, S. Bhargava, and K. Mohan, "Optimal stacked sparse autoencoder based traffic flow

prediction in intelligent transportation systems," *Virtual and Augmented Reality for Automobile Industry: Innovation Vision and Applications*, vol. 412, pp. 111–127, 2022.

[33] A. Muneer and S. M. Fati, "A comparative analysis of machine learning techniques for cyberbullying detection on twitter," *Future Internet*, vol. 12, no. 11, p. 187, 2020.

[34] R. Ram Bhukya, B. Hardas, T. Anil Kumar, and M. Ashok, "An automated word embedding with parameter tuned model for web crawling," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1617–1632, 2022.

[35] H. Singh, D. Ramya, R. Saravanakumar et al., "Artificial intelligence based quality of transmission predictive model for cognitive optical networks," *Optik*, vol. 257, Article ID 168789, 2022.

Review Article

A Complete Process of Text Classification System Using State-of-the-Art NLP Models

Varun Dogra [ID],[1] **Sahil Verma** [ID],[2,3] **Kavita** [ID],[2,4] **Pushpita Chatterjee** [ID],[5] **Jana Shafi** [ID],[6] **Jaeyoung Choi** [ID],[7] **and Muhammad Fazal Ijaz** [ID][8]

[1]*School of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India*
[2]*Department of Computer Science and Engineering, Chandigarh University, Mohali 140413, India*
[3]*Bio and Health Informatics Research Lab, Chandigarh University, Mohali 140413, India*
[4]*Machine Learning and Data Science Research Lab, Chandigarh University, Mohali 140413, India*
[5]*Tennessee State University, Nashville, TN, USA*
[6]*Department of Computer Science, College of Arts and Science, Prince Sattam Bin Abdul Aziz University, Wadi Ad-Dwasir 11991, Saudi Arabia*
[7]*School of Computing, Gachon University, Seongnam-si 13120, Republic of Korea*
[8]*Department of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Republic of Korea*

Correspondence should be addressed to Jaeyoung Choi; jychoi19@gachon.ac.kr and Muhammad Fazal Ijaz; fazal@sejong.ac.kr

With the rapid advancement of information technology, online information has been exponentially growing day by day, especially in the form of text documents such as news events, company reports, reviews on products, stocks-related reports, medical reports, tweets, and so on. Due to this, online monitoring and text mining has become a prominent task. During the past decade, significant efforts have been made on mining text documents using machine and deep learning models such as supervised, semisupervised, and unsupervised. Our area of the discussion covers state-of-the-art learning models for text mining or solving various challenging NLP (natural language processing) problems using the classification of texts. This paper summarizes several machine learning and deep learning algorithms used in text classification with their advantages and shortcomings. This paper would also help the readers understand various subtasks, along with old and recent literature, required during the process of text classification. We believe that readers would be able to find scope for further improvements in the area of text classification or to propose new techniques of text classification applicable in any domain of their interest.

1. Introduction

In recent years, we have seen a growth in the amount of digital textual data available, which has generated new perspectives and so created new areas of research. With the emergence of information technology, the monitoring of such digital textual data is of great importance in many areas such as the stock market: gathering data from news sources to forecast the movement of underlying asset volatility [1], forecasting the stock prices of green firms in emerging markets [2], understanding the impact of tone of communications on stock prices [3], and determining indicators for

stock prices volatility [4]; healthcare: disease surveillance [5, 6]; politics: developing a probabilistic framework on politics using short text classification [7]; education: understanding pedagogical aspects of the learners [8]; tourism: analyzing travelers sentiments [9]; and e-commerce: predicting success by evaluating users' reviews [10].

News is widely available in electronic format on the World Wide Web these days, and it has proven to be a valuable data source [11]. The volume of news, on the other hand, is enormous, and it is unclear how to use it most efficiently for domain-specific research. Therefore, a framework or architecture is required for a domain-specific

news monitoring system, as well as a classification mechanism for classifying relevant online news into distinct subject groups automatically [5]. News monitoring is a type of oversight system that monitors and ensures the quality of each news instance generated, used, and retained for a purpose. Processes for assessing news to guarantee its completeness, consistency, and correctness, as well as security and validity, are included in these methods. There is always a need for a methodology that can extract meaningful information from a pool of textual documents belonging to distinct subject groups intended for certain research as shown in Figure 1.

Indeed, the majority of the digital data are available in the form of text, but this is usually unstructured or semistructured [12]. Thus, to make data useful for decision-making, structuring this textual data became a necessity [13, 14]. However, because of the high volume of data, it is quite impossible to process the data manually. Text classification has evolved due to this challenge. It is defined as assigning the text documents to one or more categories (called labels) according to their content and semantics. Traditionally, the majority of classification tasks were used solved manually, but it was expensive to scale. Classification can be thought of as writing rules for assigning a class to similar text documents. These rules include some related information that identifies a class. Handwritten rules can be performed well, but creating and maintaining them over time requires much manpower. A technical expert can frame rules by writing regular expressions that could maximize the accuracy of the classifier. The existing studies have proposed various techniques to automatically classify text documents using machine learning [15, 16]. In this approach, the set of rules or criteria for selecting a classifier is learned automatically from the training data. Under each class, it requires a lot of training documents and expertise to label the documents. The labeling is a process of assigning each document to its associated class. The labeling process was easier than writing handcrafted rules. Moreover, there exist variously supervised and semisupervised learning techniques that can even reduce the burden of manual labeling [17, 18]. This can be performed using automatic labeling. Automated text classification methods can be divided into three groups: rule-based methods, data-driven methods, and hybrid methods.

Using a set of predefined rules, rule-based techniques classify text into various categories as shown in Figure 2. For example, the "fruit" label is applied to any document with the words "apple," "grapes," or "orange." These techniques require a thorough knowledge of the domain, and it is difficult to maintain the systems. Data-driven methods, on the other hand, learn to make classifications based on previous data values. A machine learning algorithm can learn the inherent associations between pieces of text and their labels using prelabeled examples as training data. It can detect hidden patterns in the data, is more flexible, and can be applied to different tasks. As the title indicates, hybrid approaches use a mixture of rule-based and machine learning methods (data-driven) for making predictions.
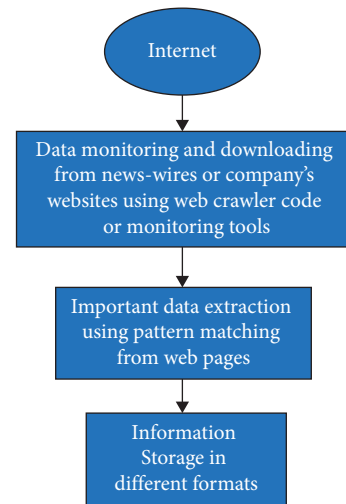
Figure 1: Monitoring and downloading relevant text documents to subject groups.
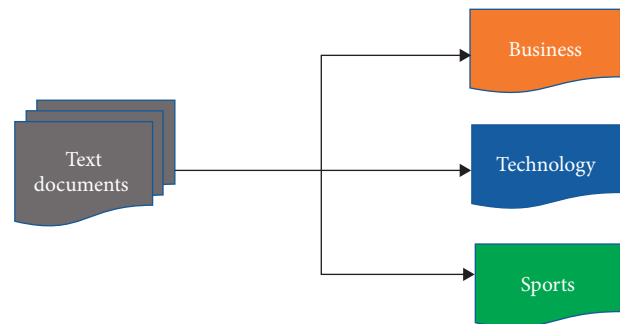
Figure 2: Labeling text documents with appropriate predefined classes or labels during the process of text classification.

In recent decades, models of machine learning have attracted a lot of interest [19, 20]. Most conventional models based on machine learning follow the common two-step method, where certain features are extracted from the text documents in the first step, and those features are fed to a classifier in the second step to make a prediction. The popular feature representation models are BOW (bag-of-words), TF-IDF (term frequency-inverse document frequency), and so on. And the common classifiers are naïve Bayes, KNN, SVM, decision trees, random forests, and so on. These models are discussed in detail in the following sections. Deep learning models have been applied to a wide variety of tasks in NLP, improving language modeling for more extended context [21–23]. These models are attempting, in an end-to-end fashion, to learn the feature representations and perform classification. They not only have the potential to uncover latent trends in data but also are far more transferable from one project to another. Quite significantly, in recent years, these models have become the mainstream paradigm for the various tasks of text classification. The following are some of the natural language challenges solved with the text classification.

Topic modeling is widely used to extract semantic information from text data. An unsupervised method of topic modeling learns the collection of underlying themes for a batch of documents as well as the affinities of each document to these topics.

News classification: online news reporting is one of the most significant sources of information. The task of finding and deriving structured information about news events in any text and assigning the relevant label is referred to as news classification.

Sentiment classification is an automatic technique of discovering views in text and classifying them as negative, positive, or neutral based on the emotions expressed in text. Sentiment classification, which uses NLP to evaluate subjective data, can help understand how people think about company's products or services.

Question answering has rapidly evolved as an NLP challenge that promises to deliver more intuitive means of knowledge acquisition. In contrast to the typical information retrieval approach of creating queries and perusing results, a question answering system simply takes user information requests stated in ordinary language and returns with a brief answer.

Language translation models have been attempting to translate a statement from one language to another resulting in perplexing and offensively inaccurate results. NLP algorithms through text classification may be trained on texts in a variety of languages, allowing them to create the equivalent meaning in another language. This approach is even applicable to languages such as Russian and Chinese, which have historically been more difficult to translate due to differences in alphabet structure and the use of characters rather than letters, respectively.

Nevertheless, it is observed that most text classification literature studies for solving NLP challenges are limited to showcasing the results of text classification using standard or state-of-the-art methods and focusing on specific research domains. For example, the authors mention the application of text analytics in the industry, but the task of monitoring and collecting text data was not detailed, and the scope of the proposed models appeared limited to particular domains [24]. In another study, the authors discuss the information extraction from tweets for monitoring trucks fleets to model truck trips, but it does not cover the feature selection or extraction methods to achieve information extraction [25]. Other studies [26, 27] focus on text classification for domain-specific search engine based on rule-based annotated data; however, it does not cover the semisupervised or unsupervised approaches of labeling data to achieve text classification [28]. Moreover, these works do not reveal the latest techniques being used in the area of natural language processing. The deep learning-based pretrained language representation model can be explored in information extraction and classification. These studies also lack in detailing the subtasks require to initiate the research in text classification, that is, data collection, data preprocessing, and semisupervised or unsupervised data labeling for training machine learning models. To the best of our knowledge, there are no similar

review studies available that cover in-depth presentations of various subtasks of text classification.

In this paper, we focus to overcome the above-mentioned issues. We put a lot of effort to create qualitative research for text classification to help us understand its subtasks or elements. Moreover, this paper presents the old and latest techniques used in each subtask of text classification as shown in Figure 3 along with their benefits and limitations. It also presents the research gap in the area of text classification by examining various existing studies. The key contribution of the study is mentioned below:

(i) Discussing the subtasks of text classification

(ii) Presenting the most recent and former techniques used in each subtask

(iii) Presenting benefits and limitations of various models used in the process of text classification

(iv) Presenting the research scope for further improvements in existing techniques and proposing new techniques with their application in different domains

Section 2 presents the process of text classification along with the comprehensive literature on each subtask; Section 3 presents the evaluation methods of classification techniques; Section 4 presents the comparison of approaches or models used in the subtasks of the text classification system mentioning their benefits and limitations; Section 5 presents the research gap and further scope for research; and Section 6 concludes the existing studies.

2. Text Classification: Framework

Text classification is a problem formulated as a learning process where a classifier is used to train to differentiate between predefined classes based on features extracted from the collection of text documents [29]. The accuracy of the classifier depends upon the classification granularity and how well separated are the training documents among classes [30, 31]. In text classification, a set of labels or classes are given, and we need to evaluate which class/label a particular text document relates to. Usually, a class or label is a general topic such as sports or business. But it may be significantly more difficult to distinguish between documents that are about more similar classes such as networks and the internet of things. Certain features represent the potential overlap between classes; the learning task would be simplified by removing such overlapping features. If the gap between classes could be increased, the classification performance would increase. This can be achieved through features weighting and selecting valuable features. Text classification has been studied and applied by many researchers in real-world scenarios such as sentiment classification of stock market news and its impact [31], news classification for syndromic surveillance [5], microblog topic classification [32], domain adaptation for sentiment classification [33, 34], and brand promotion based on social media sentiments [35, 36].
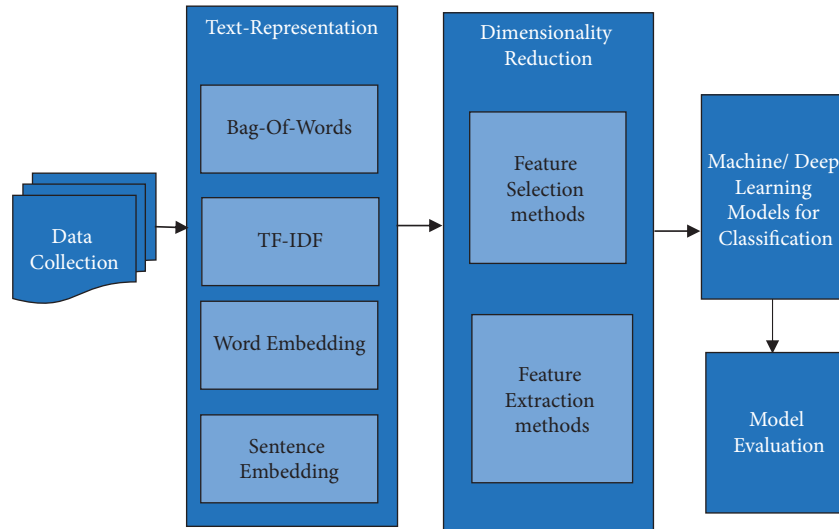
FIGURE 3: Subtasks of the text classification process cover state-of-the-art data collection, text representation, dimensionality reduction, and machine learning models for classifying text documents to an associated predefined class/label.

The text classification process is described as classifying a set of N documents; first, we build a classifier T. There is a collection of text documents D, and every text document is given a class/label by an expert. Secondly, we need to train a classifier for each class/label by giving input as a corresponding set of documents in D. Now we need to apply trained classifier C to classify N documents. We will get each document in N assigned to a predefined class/label by C. Text classification is a comprehensive process that includes not just model training but also several other steps including data preprocessing, transformation, and dimensionality reduction. The process starts with the collection of textual content from various sources. The textual content may be belonging to a domain(s) representing some events, business processes, or public information. Then these text documents require preprocessing to generate appropriate text representation for the learning model. This is done in two phases: in phase 1, the features are extracted from the processed text using any feature extraction algorithm, and in phase 2, the features are reduced by applying feature selection techniques. This reduction of features tends to decrease the dimensions of data required for the learning method. After these phases, the learning algorithms are chosen to train on data to generate the best classifier for recognizing a target category or class. This text data required to train a classifier is known as training data. The data is divided into two sets: the majority of data are taken for the training model, and the rest part of the data is taken for testing the classifier, known as testing data. Similarly, the model is trained to recognize each target class representing its data available in the associated text documents. During the testing phase, when a classification method is developed, it is executed on test data to define the target class of input text, and the result is produced in the form of weights or probabilities. Finally, the result is evaluated for its accuracy, of text classifier, using evaluation techniques. These are the main phases or subtasks of the text classification process, also shown in Figure 4. The

different approaches have been used in each phase of text classification discussed in the next subsections of the study.

2.1. Data Collection. The initial stage in text classification is to acquire text data from different sources as per the research domain. There are several online open data sets available, for example, various newsgroups (Bloomberg, Reuters, Financial Express), Kaggle, and WebKB for solving a classification problem. Researchers have used such database architecture for their research purposes [37–39]. The corpus can also be built with data that could be anything from emails, language articles, company's financial reports, medical reports, to news events. In the study, the authors have created a fine-grained sentiment analysis corpus for annotating product reviews. However, they faced the most challenging tasks that had not been targeted in applications such as sentiment analysis, target-aspect pair extraction, and implicit polarity recognition, for recognizing aspects and searching polarity with nonsentiment sentences [40].

2.2. Text Document Representation: Features Construction and Weighting. Text classification is the most demanding area of machine learning for understanding texts written in any natural language. One of the most essential tasks that must be completed before any classification process is text representation. Moreover, the texts cannot be provided as input to the machine learning models because almost all algorithms take input in numbers as feature vectors with a predefined size instead of the textual data with variable length. To resolve this issue, first textual data need to be transformed to document vectors. This can be done in two different ways in general. The first is a context-independent approach in which a document is represented as a set of terms with their corresponding frequency in the document, but they are independent of the sequence of terms in the collection. The second approach is to represent text as strings, with each
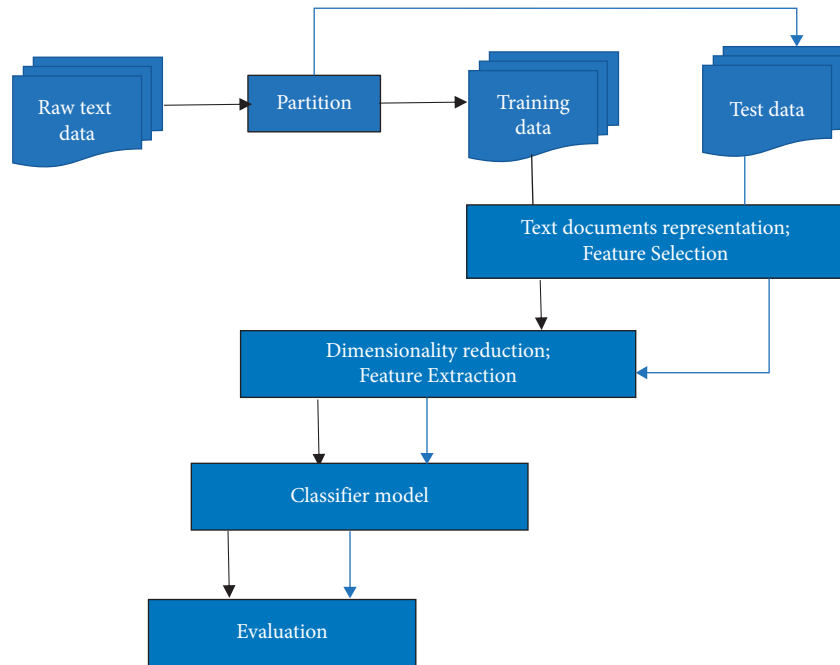
FIGURE 4: A text classification framework. Note: Black connecting lines represent training and blue connecting lines represent the testing phase.

document consisting of a sequence of terms. The following subtopic covers the various representations in natural language processing from the early days to the latest state-of-the-art models.

2.2.1. Context-Independent Approaches. The bag-of-words [41] is the most commonly used model in document transformation that considers every word in the text document as a token/feature although words' order and their context are ignored. Each word, sometimes tens or hundreds of dimensions, is represented by a real-valued vector called a one-hot representation [42]. The feature vector has the same length as the vocabulary size, and only one dimension is on as shown in Figure 5. However, the one-hot representation of a word suffers from data sparsity. On the other hand, the words with very high frequency may cause biases and dominate results in the model [44]. To overcome the weaknesses of BOW, the text documents are represented with weighted frequencies, a document-term matrix, where a column signifies a token and a row signifies a document. This scheme of assigning weights to token frequencies in the form of a matrix is called TF-IDF (term frequency-inverse document frequency). During the implementation of this model using a matrix, the value w_{ij} in each cell corresponds to the weight of t_j in d_i that is calculated as $tf\ \square t_j, d_i\ |\ n_{d_i}\square$, where $tf\ \square t_j, d_i\square$ represents the count of token t_j in text document and d_i and n_{d_i} represents the total quantity of token t_j in document d_i. Due to the simplicity of the model, this is preferably used in natural language processing. The improved features subset using this approach has been taken together with the characteristics of term frequency and document frequency [45–47]. However, even a small collection of

documents may consist of a large number of meaningful words that leads to the problem of scalability or high dimensionality. This offers opportunities to find effective ways to decrease running time or reduce high dimensionality in the case of a large number of documents.

The primary alternative has emerged in the form of statistical language modeling for modeling complex text classification or other natural language tasks. In its beginning, however, it used to struggle with the curse of dimensionality when studying typical probability functions of language models [48]. This led to the inspiration to learn distributed representations of low-dimensional space terms. The distributed representations describe a co-occurrence matrix of terms × terms that considers the frequency of each term that appears in the context of another term, with a window size of k [49]. The singular value decomposition was used for text representation, where the matrix decomposition technique was used for reducing a given matrix to its constituent matrices via an extension of the polar decomposition with the idea of making subsequent matrix calculations simpler. It gives the top rank-k constituent parts of the original data. The singular value decomposition will break this into best rank approximation capturing information from most relevant to least relevant ones [49].

The big popularization of word embedding was possibly due to the continuous bag-of-words (CBOW) paradigm to create high-quality distributed vector representations effectively. A solution is designed to counter the curse of dimensionality where a distributed representation for each word is concurrently studied along with the probability distribution for word sequences represented in terms of such representations [21]. The continuous bag-of-words is a prediction-based model that directly learns word

The quick brown fox jumped over the brown dog

	cat	the	quick	brown	Fox	jumped	over	dog	bird	flew	...	kangaroo	house
	0	1	0	0	0	0	0	0	0	0	...	0	0
	0	0	1	0	0	0	0	0	0	0	...	0	0
	0	0	0	1	0	0	0	0	0	0	...	0	0
	0	0	0	0	1	0	0	0	0	0	...	0	0
	0	0	0	0	0	1	0	0	0	0	...	0	0
	0	0	0	0	0	0	1	0	0	0	...	0	0
	0	1	0	0	0	0	0	0	0	0	...	0	0
	0	0	0	1	0	0	0	0	0	0	...	0	0
	0	0	0	0	0	0	0	1	0	0	...	0	0

time

Dictionary Size
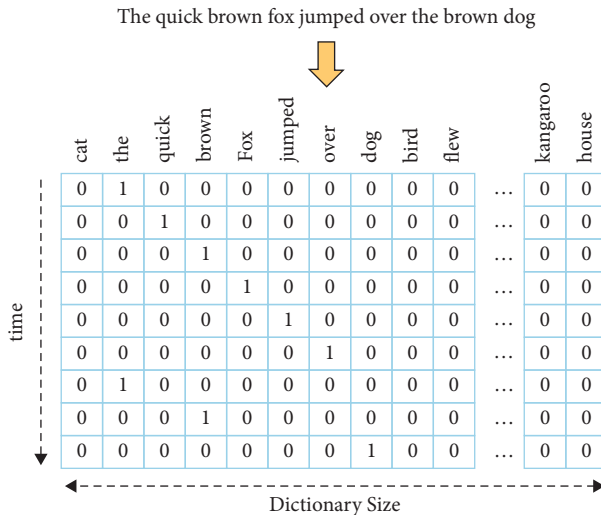
Figure 5: One-hot representation, a tensor that is used to represent each document. Each document tensor is made up of a potentially lengthy sequence of 0/1 vectors, resulting in a massive and sparse representation of the document corpus [43].

representation as shown in Figure 6(a). The distributed representations of context (or surrounding words) are combined in the CBOW model to predict the word in the middle. The CBOW has reshaped the word embedding [51]. The continuous bag of representation is applied with a neural network model to achieve improved accuracy in classification [52, 53]. Another model is designed called skip-gram that further reshaped the word embedding [54]; its architecture works in reverse of what the continuous bag-of-words model does. The model predicts each context word from the target word as shown in Figure 6(b). It iterates on the words of each sentence in the given corpus and uses the current word to predict its neighbors (its context); thus, the model is called "skip-gram" (local context window) [55].

Weighted words calculate document similarity directly from the word-count space, which takes longer to compute for large vocabularies. While counts of unique words give independent evidence of similarity, semantic similarities between words are not taken into consideration. Word embedding techniques solve this problem, but they are constrained by the need for a large corpus of text data sets for training. The word embedding algorithms were developed using the word and its closest neighbor. There was an approach suggested by authors for generating a word embedding GloVe (global vector, combines count- and predict-based methods) model for distributed word representation. The unsupervised learning algorithm, where a model is trained on overall statistics of word-word co-occurrence that how often it appears in a corpus, and the result obtains the vector representation of words with linear substructures of the word vector space [56]. GloVe's solution is to count how many times a term i (context word) in another term j (target word) occurs. The purpose is to establish a meaning for the word i and word j as to whether the two words occur close to N-word apart or not. The encoding vector includes the ratio of two words specifically recognized as a count-based system

of co-occurrence probabilities. The prediction-based approach receives popularity, but GloVe's authors claim that the count-based methodology incorporates the global statistics and may be more effective because it outperforms word representation testing on word comparison, term similarity, and called entity recognition tasks.

The enhanced text document representation system was developed to work on the issues of traditional feature-based extraction techniques that included only nouns and nouns phrases to represent the important events called event detection techniques [57]. This technique has used fewer tokens or features than bag-of-words to handle the problem of scalability or high dimensionality of documents. Furthermore, this technique has led to another representation based on named entities. The authors have presented the classification of tweets using named entity recognition to filter out noiseless required information [30, 58]. It works by finding proper nouns in the documents that belong to predefined categories. This process involves systematically assigning categories to each term or entity while developing a corpus or labeling process. But this corpus-based representation was unable to represent some domain-specific words that are infrequent during training. The authors have proposed techniques to deal with infrequent or unseen words during labeling [59].

Previous representations were not considering the morphological relation of words to disambiguate the unseen words. Many studies have presented methods that automatically extract features from the documents. These have used infrequent words that produce a high variety of low anticipated relations between the text documents. This kind of information once aggregated provides potentially less obvious and hidden relations in the text documents. Using less-frequent words with lexical constraints has reduced the associated cost of knowledge re-engineering, and it was able to process many documents from a large number of domains [59, 60]. These methods help for better representations of text documents especially handling unseen or less-frequent words. And the problem of scalability was also controlled and associated with word's semantical approach. There is another approach that was proven most efficient in a domain-specific text representation, *proper nouns*, an intermediate solution between noun phrases and named entities. This technique has reduced the ambiguity that occurred due to particular associated nouns with more than one named entity category [31]. Recent approaches are concentrating on capturing context beyond the word level to produce performance by giving a more structured and semantic notion of text [61].

2.2.2. Context-Aware Approaches. Context-aware classification approaches essentially find and employ term association information to increase classification effectiveness. They allow the presence or absence of a term to impact how it contributes to a classification outcome. Context is a concise term referring to high-level semantics. It may be taken in several ways and used in a variety of dimensions. We categorize context-based classification systems according to how
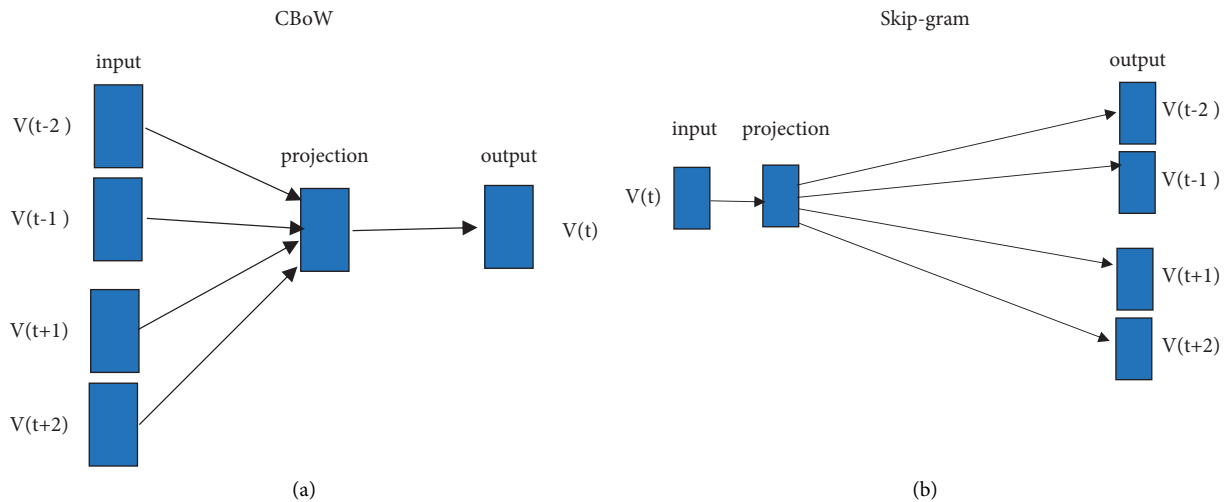
FIGURE 6: The word2vec algorithm uses two alternative methods: (a) continuous bag of words (CBoW) and (b) skip-gram (SG) [50].

the context was understood and what features were used to determine it.

The authors have come up with improved embedding for texts, such as *word2vec*, which transforms a word into an *n*-dimensional vector. To map the words into an Euclidean space, we can go through an approach to creating sequence embedding that brings a sequence into an Euclidean space. One of the sequential data function learning problems is called sequence embedding [62], where the aim is to convert a sequence into a fixed-length embedding. This approach is a highly strong tool for identifying correlations in text corpora as well as word similarity. However, it falls limited when it comes to capturing out-of-vocabulary words from a corpus. *RNN*-based models interpret the text as a sequence of words and are intended for text classification to capture word dependencies and text structures [63]. By adding a memory cell to remember values over arbitrary time intervals and three gates (input gate, output gate, and forget gate) to control the flow of information into and out of the cell, *LSTM* addresses gradient vanishing or exploding problems experienced by the RNNs. In a recursive method, the authors expand the chain-structured LSTM to tree structures, using a memory cell to store the background of multiple child cells or multiple descendants. They claim that the new model offers a realistic way to understand contact between hierarchies between long distances, for example, language parse structures [64]. To capture text features, the authors also incorporate a *bidirectional-LSTM* (Bi-LSTM) model with two-dimensional max-pooling [65]. The seq2seq model is used in various NLP applications [66, 67]. Most real-world problems have a data set with a substantial number of unusual words. The embeddings learned from these data sets are unable to produce the correct representation of the word. To do this, the data set needs to have a large vocabulary. Words that appear frequently help you create a large vocabulary. Second, when learning embeddings from scratch, the number of trainable parameters grows. As a result, the training process is slowed. Learning embeddings from scratch may also leave you confused about how the words are represented [68].

Pretrained word embeddings are the solution to all of the above difficulties. The studies have consistently shown the importance of transfer learning by pretraining a neural network model on an established problem in the field of computer vision and then doing fine-tuning utilizing the learned neural network as the foundation for a new purpose-specific model. It is demonstrated in recent years that a related approach may be effective in several tasks relating to natural language. This is another kind of word embedding; the classification algorithms provide a greater sense of learning the features of such embeddings. Yet such embeddings do not take the word order or the word meaning of each sentence into consideration. This is where *ELMo* (embedding from models of language) comes into action. ELMo is a contextual embedding that takes into consideration the terms that surround it. It models word use characteristics such as morphology and how it is used in different contexts. The term vectors are learned features of a deep bidirectional language model (biLM) internal state that is pretrained on a broad text corpus. The authors have demonstrated that these representations can be readily applied to current frameworks and greatly strengthen the state-of-the-art NLP issues such as addressing queries, textual entailment, and interpretation of emotions [62]. A Transformer [69] is another solution to working with long dependencies such as LSTM. LSTM is long short-term memory, a sort of neural network that has a "memory unit" capable of maintaining knowledge in memory over strong periods helping it to learn longer-term dependencies [22]. The Transformer is based through the encoder and decoder on an attention process as shown in Figure 7. The Transformer allows the use of this method to store knowledge about the specific meaning of a given term in its word vector.

Unlike past deep contextualized language representation studies [62] that take into account the backward model as in the ELMo bidirectional LSTM, the authors proposed a new type of language representation named *BERT*, which stands for bidirectional transformer encoder representations. BERT uses a Transformer, a mechanism of *attention* that learns
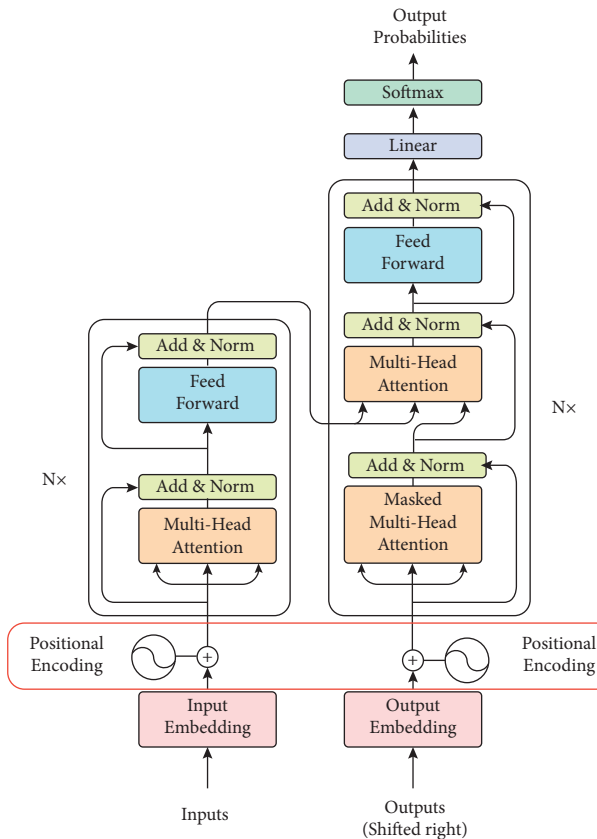
Figure 7: The transformer architecture.

contextual connections between words (or subwords) in a document. The authors have argued that conventional technologies limit the power of the pretrained representations, especially for the approaches to fine-tuning [70]. The main constraint is the unidirectional existence of modern language models, which restricts the range of frameworks that can be used during pretraining. BERT is structured to pretrain profound bidirectional representations from unlabeled text documents by jointly conditioning across both layers in both the left and right contexts [71]. In setting language modeling, transformers can acquire longer-term dependence but are constrained by a fixed-length context. The authors suggested a novel Transformer-XL neural architecture that allows learning dependence to interrupt temporal coherence beyond a fixed length. It consists of a recurrence function at the segment level and a novel positional encoding scheme [69]. Learning regarding the inductive transfer has significantly impacted computer vision, but current techniques in NLP also need complex task modifications and preparation from scratch. The authors suggested universal language model fine-tuning (ULMFIT), an important transfer learning approach that can be extended to any NLP task, and implemented techniques that are essential to fine-tuning a model of language [72].

A total of 15% of the words in every sequence are substituted with a mask token before feeding word sequences into BERT. The model then tries to determine the actual context of the masked words in the list, based on the

context given by the other, unmasked, words. The BERT loss function only considers the estimation of the masked terms and excludes the estimation of the unmasked phrases. As a result, the model converges slower than directional ones, a feature offset by its enhanced understanding of the context. Centered on BERT's masking technique [71], the authors developed a novel language representation model enhanced by knowledge-masking technique named *ERNIE* (enhanced representation by knowledge integration) [73], which involves masking at the entity level and masking at the phrase level. Strategy at the entity level covers entities that typically consist of several terms. The phrase-level technique covers the whole phrase consisting of many words that serve as a cohesive entity together. Their experimental findings indicate that ERNIE outperforms other standard approaches by obtaining modern state-of-the-art outcomes on natural language processing activities, including natural language inference, conceptual similarity, named-entity identification, emotion analysis, and question answering. DistilBERT, a technique for pretraining a smaller general-purpose language representation model that can later be fine-tuned with high performance on a wide range of tasks like its bigger equivalents, is created. While most previous research focused on using distillation to build task-specific models [74, 75], this study uses knowledge distillation during the pretraining phase and demonstrates that it is possible to reduce the size of a BERT model by 40% while retaining 97% of its language understanding capabilities and being 60% faster [76].

The Transformer paradigm is generally popular in several tasks relating to natural language processing. Using a transformer is therefore also an expensive operation since it requires the method of self-attention. The Transformer employs an encoder-decoder design that includes stacked encoder and decoder layers. Two sublayers comprise encoder layers: self-attention and a positionwise feed-forward layer. Self-attention, encoder-decoder attention, and a positionwise feed-forward layer are the three sublayers that comprise decoder layers. Self-attention assumes we are conducting the task of attention on the sentence itself, as compared to two separate sentences. Self-attention allows defining the connection in a single sentence between the words. It is the function of self-attention that adds to the expense of utilizing a transformer. The quadratic structure of self-attention, however, constrains its operation on long text. Attention is described using the (query, key, and value) model. A query Q is a "context," and in previous equations, the prior concealed state is employed as the query context. Based on what we already know, we want to know what happens next. The value represents the features of the input. The phrase "key" is just an encoding of the word "value." To attract attention, the query's relevancy to the keys is established. The associated values that are unrelated to the query are then hidden. The authors follow a method of fine to coarse attention on multi-scale spans by binary partitioning (BP); they suggest *BP-Transformer*. BP-Transformer has a strong balance between the complexity of computations and the capability of models. The authors performed a series of experiments on text classification and language

processing, showing that BP-Transformer performs superior to previous self-attention models for long text [77]. The Binary-Partitioning Transformer attempts to boost the self-attention mechanism's usefulness by considering the transformer as a graph neural network. Any node in this graph represents an input token.

Another research suggests a variant of the Neural Attentive Bag-of-Entities, which is a neural network algorithm that uses entities in a knowledge base to conduct text classification. Entities include unambiguous and specific syntactic and semantic signs that are useful for catching semantics in documents. The authors put together easy high-recall dictionary-based entity recognition, with a neural attention system that helps the model concentrate on a limited number of unambiguous and specific entities in a text [78]. The model first identifies entities to whom this name might be addressed (e.g., Ap Inc., Apple (food)) and then describes the entity using the weighted average of all entities' embedding. The weights are measured using a modern method of neural attention that helps the model concentrate on a specific subset of entities that are less ambiguous in context and more important to the document.

It is the time for NLP when the transition started as we listed the unsupervised pretrained language models that had made breakthroughs in different tasks of understanding the natural language, such as named-entity identification, emotion interpretation, and question-answer records for art performance beginning one after another in that short period. These NLP models indicate that a lot more is yet to come, and the authors look forward to researching and implementing them. However, the authors proved that a single pretrained language model may be applied as "a zero-shot task transfer" to execute basic NLP tasks without the requirement for fine-tuning on a training example data set. While this was an encouraging proof of concept, the best case performance only equaled certain supervised baselines on a single data set. Performance on most tasks was still well behind even simple supervised baselines. Across one order of magnitude of scaling, the study found generally consistent log-linear trends in performance on both transfer tasks and language modeling loss. GPT-3 performs well on NLP tasks in the zero- and one-shot settings and, in the few-shot setting, is sometimes comparable with, and occasionally surpasses, the state of the art (although the state-of-the-art is held by fine-tuned models). This might imply that bigger models are better meta-learners [79].

GPT-3 approaches the performance of a fine-tuned RoBERT as a baseline on the "Challenge" version of the data set, which has been filtered to questions (multiple-choice questions on common sense reasoning collected from 3rd to 9th-grade science examinations) that standard statistical or information retrieval algorithms are unable to accurately answer. GPT-3 marginally outperforms the same fine-tuned RoBERT baseline on the "Easy" version of the data set. However, both of these findings are much inferior to MBART's overall SOTAs.

2.3. Data Preprocessing: Data Cleaning. There are a lot of text data available today, and data are being grown daily in

structured, semiunstructured, or fully unstructured forms. To perform the text classification task, it is always required to process the raw corpus data. There are many steps involved in data processing; generally, data cleaning, that is, organizing the data as per the structure and removal of unneeded subtexts; tokenization, that is, breaking up text into words; normalization, that is, converting all texts into the same case, removing punctuation (stemming leaves out root forms of the verb and lemmatization); and substitution, that is, identifying candidate words for translation, performing word sense disambiguation [80]. In one of the studies [81], the researchers have also focused on how machine learning techniques are needed to design to recognize similar texts when text data are downloaded from multiple and heterogeneous resources. In the labeling task, the text documents are labeled with two commonly used approaches; one is to label each part of the text individually, and the second is to label the group of texts. The first approach includes different supervised learning methods, and the second is called multi-instance learning [82, 83].

2.4. Data Preprocessing: Dimensionality Reduction. Dimensionality reduction is a crucial approach in data preprocessing for preparing data for text classification. It is done to reduce the classifier's memory requirements and execution time, hence increasing the learning model's efficiency and efficacy. The dimensions of data are increasing as the volume of data grows. To map large dimensions to space with low dimensions, it becomes necessary to reduce the dimensions of the data [84, 85]. The purpose of decreasing high-dimensional space is to find a subdimensional space that is less complex and can adjust the learning model to the greatest extent possible. In some cases, several researchers have noticed that the number of features in samples is substantially larger than the number of samples. This leads to a problem known as overfitting [86]. As a result, dimensionality reduction becomes necessary to avoid overfitting. Feature selection and feature extraction are two significant subtasks in lowering dimensionality. The process of supplying the part of the original attributes that are necessary for the task is known as feature selection. Feature extraction is a technique for changing space with many dimensions into a new space with few dimensions, to increase data variance [87].

2.4.1. Standard Feature Selection Methods. The following are the goals of the feature selection method:

(i) To improve the predictability of the classifiers

(ii) To create a cost-effective classifier while also speeding up the process

(iii) To improve the clarity of the data-gathering approach

The approach of finding a portion of the original attributes that are significant for the training set is known as feature selection. It is used to create an effective classifier while keeping important lexical properties in vocabulary [88]. It removes the noisy attributes that tend to reduce

accuracy [89]. The goal of feature selection is to decrease the feature space by picking a subset of Kay attributes and minimizing overfitting while maintaining text classification performance [90]. For instance, a feature or attribute set $X = \{X_1, X_2, \ldots, X_N\}$

When N is a group of feature sets, $2N$ feature subsets are created, each of which is represented by a vector of size N. The methods identify a feature subset of size K, $K < N$ without losing the accuracy of the full feature set. It has been the subject of investigation, and the writers have so far provided many techniques. Filter, wrapper, and embedded are the three types of approaches available [91, 92]. Filter-based feature selection offers several ways to evaluate the information value of each feature. The filter approach picks the top-N features based on the results of various statistical tests to identify a connection with the target label or to establish which attributes are more predictive of the target label, and it is independent of any learning algorithms. Because this technique ignores feature dependencies, it is computationally efficient. The wrapper technique evaluates a subset of features based on their utility to a given class. It uses a learning model to assess the subset of features based on their predictive power, but it is significantly very costly owing to repetitive learning and cross-validation. Embedded techniques are analogous to wrapper methods, except that they include feature selection during the training phase [93].

(1) Filter methods

Univariate Feature Selection. To determine the link between the features and the target variable, univariate feature selection evaluates each feature independently. Because they pertain to linear classifiers created using single variables, the following univariate feature selection methods are linear.

The filter approach chooses attributes without focusing on the core goal of improving any classifier's performance. To score the attributes, it uses the data's most important attributes. If d features or attributes are identified as S, the goal of a filter-based approach is to pick a subset of $m < d$ features, T, which maximizes some function F:

$$\tau^* = \operatorname{argmax}_{\tau \subseteq S} F(\tau), \text{s.t.} |\tau| = m. \tag{1}$$

It finally settles on the top-m rated features with the highest scores. This number is known as joint mutual information, and maximizing it is an NP-hard optimization problem since the number of potential feature combinations rises exponentially.

The following are the often used linear univariate filter techniques in text classification:

The *information gain* approach selects features based on the item's frequency concerning the class/label prediction. Researchers have demonstrated that by removing superfluous features without modifying the features, the approach may lower the vector dimensionality of text and enhance classification results [94]. It adds the most value when the text corresponds to a

certain label or class, and the word is also present in the document. It can be written as follows:

$$IG(t) = -\sum_{i=1}^{m} P(C_i) \log p(C_i) + p(t) \sum_{i=1}^{m} P(C_i \mid t) \log p(C_i \mid t)$$

$$+ p(\bar{t}) \sum_{i=1}^{m} P(C_i \mid \bar{t}) \log p(C_i \mid \bar{t}). \tag{2}$$

The utility of feature t in the classification is measured by this formula. If *IG* is higher than the prior value without the feature t, the current feature t is more relevant for classification. In other words, the discriminating power of the term t increases as the value of the information gain *IG* increases. Here, *IG* stands for information gain; C_i is the i-th class; $P(C_i)$ is the probability of an i-th class; and m is several target classes. $P(t)$ is the probability the feature t appears in the documents and the probability $P(\bar{t})$ for feature; t does not appear in the document. $P(C_i|t)$ is the conditional probability of the feature t appearing in i-th class. $P(C_i \mid \bar{t})$ is the conditional probability of the feature t that does not appear in i-th class.

The *Chi-square test* is a statistical strategy for assessing the relationship between a set of categorical features using their frequency distribution and determining how much the findings differ from the predicted output [95, 96]. This may be determined given events A and B, which are considered to be independent if

$$p(AB) = p(A)p(B). \tag{3}$$

The occurrence of the term and the occurrence of the class are the two events in feature selection. The terms are then ranked according to the following value. Chi-square can be calculated from the following equation:

$$\text{CHI}^2(t, C) = \sum_{t \in \{0,1\}} \cdot \sum_{C \in \{0,1\}} \frac{\left(N_{t,C} - E_{t,C}\right)^2}{E_{tC}}, \tag{4}$$

where t denotes the feature, C denotes the specific class, $N_{t,C}$ is the frequency of feature t and class C occurring together, and $E_{t,C}$ is the frequency of feature t occurring without class C. The chi-square between each feature and class is computed, and the features with the highest chi scores are chosen.

Fisher score calculates the variance of the predicted value from the actual value to get the information score or how much knowledge one variable has about the unknown parameter on which the variable depends, and when the variance is the smallest, the information score is the highest. For a support-vector-based feature ranking model, researchers employed Fisher's linear discriminant [97–99].

For instance, let μ_j^k and σ_j^k be the mean and standard deviation of the k-th class, concerning the j-th feature.

Let μ^j and σ^j represents the mean and standard deviation of the entire training data concerning the j-th feature. The Fisher equation for determining the j-th feature's score is stated as follows:

$$F\left(x^j\right) = \frac{\sum_{k=1}^{c} n_k \left(\mu_k^j - \mu^j\right)^2}{\sigma_j^2}, \tag{5}$$

where σ_j^2 is computed as $\sum_{k=1}^{c} n_k (\sigma_k^j)^2$. Top-$m$ features with higher fisher scores are chosen by the algorithms.

Pearson's correlation coefficient is used to measure linear dependency between two continuous variables by dividing their co-variance by the product of their standard deviation, and its value ranges from -1 to $+1$. In two variables, a–1 value signifies a negative correlation; a +1 value shows a positive correlation; and a 0 value represents no linear association [93].

Using vectors, Pearson's coefficient r can be computed as follows:

$$r = \frac{\left(x_1 - \overline{x}_1 L\right)\left(x_2 - \overline{x}_2 L\right)}{\left(|x_1||x_2|\right)}, \tag{6}$$

where $\overline{x}_1$ is the mean of the vector x_1 and similarly for x_2, L is the vector of 1s, and $|x|$ is the magnitude of vector x.

Variance threshold is a technique for reducing vector dimensionality by deleting all low-variance features. Features that have a lower training-set variance than the threshold will be deleted [100, 101].

$$s = \frac{\sqrt{\sum \left(x_i - \overline{x}\right)^2}}{(n-1)}. \tag{7}$$

The equation may be used to find features that have a variation below a given threshold. When the feature does not vary much within itself, it is seen to have low predictive potential.

Multi-Variate Filter Methods. During the assessment of the multi-variate filter selection approach, the interdependencies of features are also taken into account to choose relevant features.

It is based on mutual information that discovers the features in a feature set with the highest dependency with the target label. However, it is not appropriate for use when the goal is to achieve high accuracy with a small number of features.

Alternatively, it may utilize max relevance, which detects features with a dependency by averaging all mutual information values between all features x_i and target label c. S refers to features, and I represents mutual information; in the following equation, it is calculated between feature i and class c:

$$\max D\left(S, c\right), D = \frac{1}{|S|} \sum_{x_i \in S} \left(I\left(x_i; c\right)\right). \tag{8}$$

However, this results in a high level of redundancy, that is, a higher level of a dependency across features. As a result, to locate mutually exclusive features, minimum redundancy can be used [102].

$$\min R\left(S\right), R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I\left(x_i, x_j\right), \tag{9}$$

where $I(x_i, x_j)$ is the mutual information between feature i and j.

Multi-Variate Relative Discriminative Criterion. The author offers a multi-variate selection strategy that takes into account both feature relevance and redundancy in the selection process. The RDC is used to assess the relevance, whereas Pearson's correlation is used to assess redundancy between features [103]. This measure boosts the rankings of terms that are exclusively found in one class or whose term counts in one class are much higher than in the other.

$$\mathrm{RDC}\left(w_i, tc_j(w_i)\right) = \frac{\left(\left|df_{\mathrm{pos}}(w_i) - df_{\mathrm{neg}}(w_i)\right|\right)}{\min\left(df_{\mathrm{pos}}(w_i), df_{\mathrm{neg}}(w_i)\right) * tc_j(w_i)}, \tag{10}$$

where $df_{\mathrm{pos}}(w_i), df_{\mathrm{neg}}(w_i)$ are the collection of positive and negative text documents, respectively, in which the term w_i is occurred. The word may be repeated several times in specific documents and represented by $tc_j(w_i)$. Instead of adding together RDC values for all term counts of a term, the area under the curve (AUC) for a difference graph is treated as term rank.

Researchers have been looking for novel approaches to increase classification accuracy while also reducing processing time. The author has provided a differentiate feature selector, a filter-based strategy that has picked unique features that have term properties while eliminating uninformative ones [92]. It provided efficiency by reducing processing time and improving classification accuracy.

(2) Wrapper Methods. Wrappers' approaches are bound to a certain classifier; the methods choose a subset of features based on their influence on the classifier by assessing the prediction performance of all potential feature subsets in a given space. It signifies that the features subset will be assessed by interacting with the classifier, which will improve the classification technique's accuracy. As the feature space expands, the computing efficiency suffers as a result of this method. Wrappers are used to choose features for other models as filters. The procedure may be accomplished in three ways: the first methodology employs a best-first search technique; the second methodology employs a stochastic approach such as random selection; and the third

methodology use heuristics such as forward and backward passes to include and omit features.

Multi-Variate Feature Selection. Univariate feature selection approaches are computationally efficient, but they eliminate features owing to a lack of interaction between features that, when combined, may have offered important information regarding classification [104, 105]. When evaluating the performance of features, multi-variate takes into account the interdependencies between them. "Linear multi-variate" employs linear classifiers made up of a subset of features, with the score of feature subsets being calculated based on classification performance. Nonlinear multi-variate, on the other hand, use nonlinear classifiers to complete the task.

The following are the most often used *linear multi-variate wrapper* approaches in in-text classification:

Recursive Feature Elimination. It is a recursive strategy that ranks features according to a key measure. During each cycle, the significance of features is assessed, and less relevant features are removed. To design ranking, the opposite process is utilized, in which features are rejected. From this rating, this technique extracts the top-N features [106]. This is a greedy optimization that seeks the highest performing feature subset.

Forward/Backward Stepwise Selection. It is an iterative procedure that begins with the examination of each feature and picks the one that produces the best performing model, based on some predetermined criteria (like prediction accuracy). The next step is to examine every potential combination of that selected feature and the following feature, and if it improves the model, the second feature is chosen. The model continuously appends the list of features that best improve the model's performance in each iteration until the requisite features subset is picked. In the backward feature selection approach, the method starts with the whole collection of features and discards the least relevant feature in each iteration, improving the method's speed. This method is repeated until no improvement is shown when features are removed, and the best subset of features is found. In comparison to other techniques, the researcher developed a rapid forward selection methodology for picking the optimal subset of features that required less computing work [107].

The *Genetic Algorithm* uses a feature set to generate a better subset of features that are free of noise. At each step, a new subset is formed by picking individual features in the correct sequence and merging those using natural genetics procedures. The result is cross-validated variance divided by the percentage of right predictions. The end outcome may be mutated. This procedure aids in the creation of a feature set of individual features that are more appropriate for the model than the initial feature set. The chaotic genetic algorithm was designed to simplify the feature

selection procedure and improve the classification technique's accuracy [108, 109].

The commonly preferred *nonlinear multi-variate wrapper* methods in the text classification are discussed as follows:

Nonlinear kernel multiplicative updates entail iteratively training a classifier and rescaling the feature set by multiplying it by a scaling factor that lowers the value of less impacted features. Nonlinear techniques can outperform linear algorithms by selecting a subset of features [110].

Relief is based on instance-based learning. Each feature receives a value ranging from −1 to +1 based on how well it matches the desired label. The algorithm's scope is binary-classification-compatible [111, 112].

(3) Embedded Methods. In terms of computing, "embedded methods" outperform wrappers, but they conduct selection features as a subpart of the learning methodology, which is primarily exclusive to the learning model and may not function with any other classifier.

The commonly preferred embedded methods in the text classification are discussed as follows:

In social sciences, the *LASSO* method is generally used [113]. To alleviate the dimensionality problem, it penalizes features with large coefficients by inserting a penalty during the log-likelihood maximization procedure. By picking a correct weight and reducing dimensionality, LASSO assigns zero to some coefficients. When there is a strong correlation between some features, it creates a difficulty [114].

Ridge Regression lowers the complexity of a model by reducing coefficients while keeping all of its features. The issue with ridge regression is that features are retained. If the feature collection is huge, the problem remains complicated [115].

Elastic Net calculates a penalty that is a mix of LASSO and ridge penalties. The elastic net penalty may be readily handled to give LASSO or ridge penalties extra power. It has a grouping effect, with high correlation features tending to be in or out of the feature subset. It incorporates both L1 and L2 regularization techniques (LASSO and ridge). By fine-tuning the settings, it aids in the implementation of both strategies [116].

2.4.2. Text Feature Extraction Methods. After selecting the features and representing N documents by d-dimensional features *vectors* $\{X_1, X_2, \ldots, X_N\}$. Sometimes original terms in the form of features may not be optimal dimensions for text document representation. The text feature extraction methods try to solve these problems by creating new feature space Y or artificial terms [117]. It requires (a) a method to convert old terms to new terms and (b) a method to convert document representation from old to new. A popular example commonly used for this purpose is principal component analysis (PCA) in which a feature set Yi is selected in a manner that the variance of the original feature vectors is maximized in the direction of new feature vectors. This is

done by computing eigenvectors of the covariance matrix of the original vectors. The drawback of the PCA is the time it takes to evaluate eigenvalue decomposition to compute the principal component for each class/label when applied to a large data set. This is overcome by the researcher by using the power factorization method (PFM) to find a fixed quantity of eigenvectors from a data set [118]. Another commonly preferred method for feature extraction is latent semantic indexing (LSI) that uses singular value decomposition of the term correlation matrix computed from a large collection of text documents. This technique is used to address the problem of deriving from the use of synonymous and polysemous words as dimensions of the text document representation. But the disadvantage is to compute the correct number of latent components that proves computationally expensive. Another method that helps for optimal discrimination of data is linear discriminant analysis (LDA) [99]. It identifies the linear collection of features that best explain the data. It tries to find the model that can explicitly differentiate between the classes of data. Latent Dirichlet allocation is another method that explains that each text document is a mixture of latent topics and each word in that document is attributable to one of the topics of that document. This is most preferred for topic modeling [119]. This is a generative probabilistic model.

A newer approach is a simplified version of stochastic neighbor embedding that creates much better visuals by eliminating the potential to cluster points together in the map's centers known as t-SNE. It visualizes high-dimensional data by assigning a two- or three-dimensional map to each data point. When it comes to constructing a single map that displays structures of several sizes, t-SNE outperforms previous approaches. On almost all of the data sets, the analysis shows that t-SNE produces visuals that are much superior to those produced by the other approaches [120]. Furthermore, the authors provide a technique UMAP (uniform manifold approximation and projection) that is comparable to t-SNE in terms of visualization quality and, in certain ways, retains more of the global structure while providing better run time efficiency [121]. It is based on Laplacian eigenmaps as a mathematical foundation. Umap can scale to far bigger data sets than t-SNE. It is a general-purpose dimension reduction strategy for machine learning since it has no computational constraints on embedding dimensions. The approach is used in the study to evaluate the uniqueness of subjects, important phrases and features, information dissemination speed, and network behaviors for COVID-19 tweets and analysis [122]. To increase the detection of relevant themes in the corpus and analyze the quality of created topics, they use UMAP, which finds distinctive clustering behavior of separate topics [120]. Another study used UMAP to depict the matching word vector spaces of a pretrained language model using LaTeX mathematical formulations. In the LaTeX formula domain, they develop a state-of-the-art BERT-based text classification model augmented by unlabeled data (UL-BERT) [123].

2.5. Classifiers for Classification Task.

The classifier is trained based on the selected features from the text documents. The selection of appropriate features in feature space decides the performance of learning models. Machines understand numbers more easily than texts as input. So texts as tokens are required to be converted into numbers (vectorization) for most of the learning algorithms. Vectors are combined to originate vector space to apply statistical methods for checking document relatedness. Each algorithm offers a different document representation for text classification. Researchers offer several classification methods that work on the vector representation of texts. The basic assumption for vector representation of texts is known as the contiguity hypothesis. It states that text documents belonging to the same class develop a contiguous region and regions of different classes do not overlap. The relatedness of the documents can be evaluated on 2D space based on cosine similarity or Euclidean distance.

The authors have presented naïve Bayes, a linear classifier, approach to vectorize the text documents according to probability distribution with two commonly used models: multi-variate Bernoulli event and multi-nomial event; the features with the highest probability were chosen to reduce the dimensionality [124].

$$PC_k|d_i = Pd_i|C_k * \frac{P(C_k)}{P(d_i)}. \quad (11)$$

The output of the classifier is the probability of the text document d_i is belonging to each class C_k, and it is a vector of C elements. In a way of text classification scenario, we could compute $Pd_i|C_k$ using bag-of-words as follows:

$$Pd_i|C_k = P(BoW(d_i)|C_k) = Pw_{1,i}w_{2,i},,,w_{|V|,t}|C_k. \quad (12)$$

The problem can be reduced to compute the probability of each word $w_{j,i}$ in class C_k as follows:

$$Pd_i|C_k = \prod_j^{|V|} Pw_{j,i}|C_k. \quad (13)$$

The naïve Bayes has a high bias for a nonlinear problem because it can model one type of class, that is, a linear hyperplane. The bias is the statistical method of evaluating the performance of a classifier that how accurately a classifier classifies the texts into the correct class with less error. The learning task is the last activity in classification, but reducing the feature dimensions is more concerned with the efficiency of the classifier model. The possibility of salient feature reduction caused by using classifiers can be overcome by using the model SVM that identifies the best decision-boundary between feature vectors of the document with their categories [125]. The SVM entails an optimal classifier that guarantees the lowest classification error. The SVM computes a hyperplane that falls between the positive and negative example of the training set.

$$D = \left\{ (x_i, y_i) \middle| x_i \in R^D, y_i \in \{-1, 1\} \middle| \right\} i =_i^n, \quad (14)$$

where the minus sign represents the negative hyperplane, the positive sign points to the positive hyperplane, i ranges from 1 to L (training examples), (x_i, y_i) represents feature vectors of each document, R^D is a vector space having a dimension of D. and D is evaluated to +1 and −1 for positive and negative hyperplanes, respectively.

The naïve Bayes itself results in the best classification model if it is trained on a high volume of data. However, feature reduction remains an issue. So naïve Bayes is used as a prestep to SVM that converts text documents into vectors before the classification task starts. This resulted in improving the whole system while spending quite an appropriate classification time by reducing to low-dimensional space. But, in certain cases, the majority of features are redundant with each other; the author has presented a divergence-based feature selection method without relying on a complex dependence model where the maximum marginal relevance-based feature selection was outperformed by the SVM [126]. The paper has suggested the need for novel criteria to measure the relevance and novelty of features separately and provided the linear combination as the metric.

The studies have mentioned the KNN, a nonlinear classifier, where the algorithm classifies the document by moving through all the training documents that are like that document. The KNN model frames the documents in the Euclidean space as points so that the distance between two-point u and v can be calculated as follows:

$$D(u, v)^2 = u - v^2 = (u - v)^T (u - v) = \sum_{i-1}^{d} (u_{i-} v_i)^2. \quad (15)$$

The classifier finds the K-value that is the factor that represents a collection of documents from all the documents closest to the selected document in that space [127]. If there are too many features, KNN may not operate effectively. As a result, dimensionality reduction techniques such as feature selection and principal component analysis may be used.

$$\hat{y}(x) = y_{n^*} \quad \text{where} \quad n^* = \arg \min_{n \in D} \text{dist}(x, x_n). \quad (16)$$

The study mentions that using KNN increases the overhead to calculate the K-value of all the documents with all other training documents with the largest similarity or closet to the selected document. Also, the variation in the number of training sample documents in different categories leads to a decline in accuracy. Due to the high variance and complex regions between classes, it becomes sensitive to noise documents. Sometimes, the document tends to misclassify if it occurs very relevant to a noise document in the training set and sometimes accurately classified if there is no presence of noise documents in the training set close to them. This ends up in high variance from the training set to the training set. High variance leads to overfitting. The goal of finding a good learning classifier is to decrease the learning error. Learning error is calculated from bias and variance or bias-variance trade-off. The traditional algorithms possess some limitations that attract the researchers to improve the efficiency by (a) reducing the computational overheads by establishing low-dimensional space, (b) speeding up the computational capacity of finding nearest neighbors in KNN or locating decision boundaries in SVM, and (c) increasing efficiency by not compromising accuracy [128]. The model is chosen that optimizes the fit to the training data.

$$H^* = \arg \max_{H} \text{fit}(H|D). \quad (17)$$

The supervised learning classification algorithms such as naïve Bayes, SVM, and KNN use a bag of words to design a classifier. None of these methods take the order of words into consideration that can lead to the loss of some valuable information. In natural-language-based problems, the order of the words, for example, multi-words, has a meaning (like names of organization or person) that is not considered by the learning models trained on individual words of the texts. The algorithm n-grams consider the sequence of n-adjacent words from the selected text phrases. The n-grams behave like the individual word's representation as feature vectors. The value of n may range from 1 to the upper value [129]. This proves very beneficial in short text documents where the number of n-grams is less in number [89]. The author proposes another representation for character n-grams to introduce the enhancement of the skip-gram model, which considers subword into account. It considers the word morphology while training the model, and words are represented by the sum of its character n-gram [55].

Many researchers have used a decision-tree-based algorithm (decision support tool) that represents a tree-structure-based graph of decisions [130]. The commonly used decision tree algorithms are ID3, C4.5, and C5. The algorithm presents each intermediate node (labeled as terms and branches represent weight) that can split into subtrees and ends at leaf nodes (represents the class/label/outcome of the problem). Decision tree structures are rule-based solutions. A rule can be designed by forming a conjunct of every test that occurs on the path between the root node and the leaf node of the tree. The rules are formed after traversing every path from a root to the leaf node. Once the decision tree and rule are framed, it helps assign the class/label for a new case [129, 131]. It was evaluated in the study that decision trees result better than naïve Bayes in terms of accuracy but a little worse than KNN methods [132]. As a result, the authors introduce the boosting data classification approach. The boosting algorithm is a method for combining many "poor" classifiers into a single, powerful classifier. The boosting technique is used on decision trees in the study, and the boosted decision tree performs better than an artificial neural network [133]. It is expected to find widespread use in a variety of fields, particularly text classification. Gradient tree boosting is another boosting strategy that builds an additive regression model using decision trees as the weak learner. Trees in stochastic gradient boosting are trained on a randomly selected portion of the training data and are less prone to overfitting than shallow decision trees [134].

Neural networks or deep learning systems use several processing layers to learn hierarchical data representations and have reached state-of-the-art outcomes in most domains. In the sense of natural language processing (NLP), several model designs and approaches have recently

progressed. In areas such as computer vision and pattern recognition, deep learning architectures and algorithms have also made remarkable progress. Recent NLP research is now primarily focused on the application of new deep learning approaches, continuing this development. In areas such as computer vision and pattern recognition, deep learning architectures and algorithms have also made remarkable progress. Recent NLP research is now primarily focused on the application of new deep learning approaches, continuing this development. The performance of word embedding and deep learning strategies referred to in the following section is driving this development.

In the late 90s, the researcher found an application of nonlinear neural networks to text classification or topic modeling [135]. In this model, a three-layered neural network was designed to learn a nonlinear mapping from training documents to each class/label. Later on, the researcher proposes convolutional neural network (CNN) for text classification by considering the order of words in the phrases, and this outperforms SVM in terms of error rate [136]. CNN uses the vector representation of text data considering the order of words. Each word is considered a pixel, and the document is treated as an image. Then the image is taken into $|D| \times 1$ pixels, and each pixel represent a word as a $|V|$ dimensional vector. For instance, vocabulary $V = \{$"classification", "course", "I", "love," "NLP'," "text"$\}$, and words are taken as a dimension of vectors in alphabetical order. And document $D = $ "I love NLP course". Then the document vector would be $X = [0010000 \mid 000100 \mid 00001 \mid 010000]$.

The researcher mentions that to reduce the dimensionality of vector space, the vocabulary size must be kept low. Also, the n-gram algorithm ignores the fact that some n-grams share the constituent words; this is overcome by CNN that learns the embedding of text regions by providing CNN with the constituent words as input, and this technique provides higher accuracy. To construct an informative latent semantic representation of the sentence for downstream activities, CNNs have the potential to extract salient n-gram features from the input sentence. The author proposes a three-way enhanced CNN for classification for sentiment analysis, where decisions are divided into three parts accept, reject, and delay. The instances in boundary regions that are neither in accept nor reject are reclassified by another classification model. This guarantees the enhancement in CNN to deal with boundary regions in a better way, resulting in model 3W–CNN [137]. Another study has shown the application of CNN in document modeling for personality detection based on text in the context of sentiment analysis [34]. Overall, in contextual windows, CNNs are highly successful in mining semantic hints. They are very data-heavy models, though. They have a huge range of parameters that are trainable and need tremendous training data. This raises a concern as data shortage happens. Another unresolved concern with CNNs is their failure to model contextual long-distance data and maintain sequential order in their representations [138].

Deep neural networks are difficult to train on data; it requires a lot of resources to get high performance. The feed-forward neural network commonly known as multi-layer perceptron (MLP) is the most preferred technique in

classification problems. In a feed-forward neural network, the information travels in one direction from the input layer to hidden layers and then followed by the output layer. They have no memory of the input received previously so lack in predicting what comes next. To overcome this, RNN (recurrent neural network) is preferred where information moves through the loop. In this paragraph, we discuss the fundamental characteristics that have favored the popularization of RNNs in a variety of NLP tasks. Since an RNN performs sequential processing in sequence by modeling units, it may have the ability to generate the intrinsic sequential structure present in language, where characters, words, or even phrases are units. Based on the previous words in the sentence, words in a language establish their semantic meaning. The disparity in interpretation between "computer" and "computer vision" is a clear example that states this. RNNs are perfect for language and related sequence modeling activities to predict certain context dependencies, which turned out to be a clear incentive for researchers to use RNNs over CNNs in these fields.

RNNs were originally three-layer networks in the NLP sense [139]. While deciding on the current input layer, it considers what it has learned from the previous inputs. Basically, in the architecture of simple RNN, the hidden units create internal representations for the input patterns and recode these patterns in feed-forward networks using hidden units and a learning algorithm in a way that allows the network to generate the appropriate output for a given input. Typically, the hidden state of the RNN is assumed to be the most important feature. It can be regarded as the memory portion of the network that accumulates data from other steps.

The formula of the current state in RNN can be written as mentioned below. A nonlinear transformation such as tanh, or ReLU, is taken to be the function f.

$$h_t = f\left(h_{t-1}, x_t\right), \tag{18}$$

where h_t is the new state, h_{t-1} is the previous state, and X_t is the input at time t.

The tanh function is commonly used as an activation function. The weights can be defined as the matrix W_{hh}, and input is defined by the matrix W_{xh}:

$$h_t = \tan h\left(W_{hh}h_{t-1} + W_{xh}X_t\right). \tag{19}$$

The output can be calculated during test time as follows:

$$y_t = W_{hy}h_t. \tag{20}$$

The output is then compared to the actual output, and then the error value is computed. The network learns by backpropagating the error through the network to update the weights. But usual RNN has a short-term memory. These basic RNN networks suffer from the issue of vanishing gradient, which makes it very difficult to understand and adjust the parameters of the previous layers in the network. It is used in combination with LSTM, which has long-term memory, and it gives an extension to the memory of the usual RNN. Over the basic RNN, LSTM has additional "forget" gates, allowing the error to backpropagate over an

infinite amount of time steps. Comprising three gates: input, forget, and output gates, taking a combination of these three gates, it determines the hidden state [22]. The applications based on RNN and LSTM have been used in solving many NLP problems due to their capacity of capturing complex patterns within the text [140]. It has also been used in sequence labeling tasks in POS (part of speech) activity. It is preferably used in topic modeling for fake news [141, 142], sentiment analysis [143, 144], and negative speech detection on social media. More recently, authors have suggested another type of recurrent unit, which they refer to as a gated recurrent unit (GRU) [145]. It has been shown that RNNs employing any of these recurrent units perform well in tasks (such as machine translation, speech recognition, or depending parsing in text documents for NER) requiring long-term dependency capture. The application of gated RNN is not limited to the mentioned tasks, but it can be applied to different NLP challenges [146]. GRU is a form of recurrent neural network (RNN) that can process sequential data using its recurrent architecture. The fundamental issue in text classification is how to improve classification accuracy, and the sparsity of data, as well as semantics sensitivity to context, frequently impedes text classification performance. The study introduces a unified framework to evaluate the impacts of word embedding and the gated recurrent unit (GRU) for text classification to overcome the flaw [147].

Recurrent neural networks, in particular long short-term memory [22], and gated recurrent neural networks [62] are firmly known as state-of-the-art approaches in sequence modeling. Within training examples, the inherently sequential nature of recurring models prevents parallelization, which becomes important at longer sequence lengths, as memory limitations restrict batching through examples. Recent work, through factorization tricks [148] and conditional computation [149], has achieved substantial improvements in computational efficiency while also improving model output in the case of the latter. The authors have presented two simple ways of reducing the number of parameters and speeding up the training of large long short-term memory networks: the first is the "matrix factorization by design" of the LSTM matrix into two smaller matrices, and the second is the division into separate classes of the LSTM matrix, its inputs, and its states. However, the essential restriction of sequential computation persists.

Attention mechanisms have become an integral part of sequence modeling in different applications, allowing dependencies to be modeled regardless of their gap in the input or output sequences. The following paragraph examines some of the most prominent models of attention that have created new state-of-the-art tasks for text classification. The study's conclusion was largely focused on text classification experiments that could not be extended to many other NLP tasks [150]. The authors mentioned that attention offers a reliable explanation for model predictions, expecting these properties to hold (a) attention weights should align with feature-relevant measurements (e.g., gradient-based measurements) and (b) alternative (or counterfactual) weight configurations should result in corresponding prediction changes. They stated that in the sense of text classification,

neither property is consistently observed by a Bi-LSTM with a standard attention mechanism. The layers of attention specifically weigh the representations of the input elements; it is also often believed that attention can be used to classify information that was considered relevant by models. In another study, the authors evaluate if that assumption holds by modifying weights of attention in already trained text classification methods and examining the resulting variations in their predictions. Although experimenting with text classification, the authors note several cases in which higher attention weights correlate with a greater impact on model predictions, they also notice several cases this does not hold, that is, where gradient-based attention weight rankings predict their effects better than their magnitudes [151]. In contrast to the current work on interpretability, the authors in another research reported that they examined the attention mechanism on a more diverse set of NLP tasks that included text classification, pairwise text classification, and tasks for generating text such as neural machine translation [152].

Although the hidden vectors represented by an attention model through encoding can be interpreted as internal memory entries for the model, memory-augmented networks integrate neural networks with an external memory type that the model could learn from it and respond to. For text classification, the study proposes a memory-augmented neural network called the neural semantic encoder [153]. In another research, the authors introduce neural network architecture, the dynamic memory network (DMN), which processes input sequences and questions, shapes episodic memories, and generates appropriate responses. The model possesses an iterative mechanism of attention that allows the model to condition its attention on the inputs and outcomes of previous iterations. In a hierarchical recurrent sequence model, these outcomes are then reasoned over to produce answers [154]. The authors also stated that it is possible to train the DMN end-to-end and obtain state-of-the-art text classification results using the Stanford Sentiment Treebank data collection.

Sequential processing of text is one of the system inefficiencies experienced by RNNs. Transformers address this constraint by applying self-attention to measure an "attention ranking" in parallel for each word in a phrase or document to model the impact each word has on another. Because of this feature, Transformers allow for far more parallelization than CNNs and RNNs, allowing very large models to be efficiently trained on large quantities of data on GPU stacks [155]. The Transformer architecture is especially suitable for pretraining large text corpus, leading to significant accuracy improvements in downstream tasks, including text classification [69]. They propose a novel Transformer-XL neural architecture that allows learning dependence to interrupt temporal coherence beyond a fixed length. We have seen the emergence of several large-scale transformer-based pretrained language models in the current scenario. As stated in Section 2.2.2, these Transformers are pretrained to learn contextual text representations in much greater volumes of text corpora by predicting terms that are trained on their context. These pretrained models

were fine-tuned using task-specific tags and in many subsequent NLP tasks, especially text classification produced new state-of-the-art. Fine-tuning is supervised learning, while pretraining is unsupervised.

The authors design the largest model, OpenGPT, 1.5 B parameter Transformer that ensures state-of-the-art results and comprises 12 layers of Transformer frames, each composed of a masked multi-head attention unit, followed by a standardization layer and a forward feed layer in place. With the addition of task-specific linear classifiers and fine-tuning with task-specific tags, OpenGPT can be extended to the text classification. Unlike OpenGPT that predicts words based on previous predictions, there is another model that comes into use, that is, BERT, intended to pretrain deep bidirectional representations from the unlabeled text by conditioning in all layers on both the left and right context together [71]. For text classification, BERT variants have been fine-tuned [156]. ALBERT decreases memory usage and improves BERT's training speed [157]. Another variant of BERT, SpanBERT [158], is a pretraining method designed to accurately represent and forecast spans of text. It improves BERT by (1) masking consecutive random spans, rather than random tokens, and (2) training the span boundary representations to estimate the entire content of the masked span without relying on the individual token representations within it. Deep learning provides a way to manage massive volumes of processing and data with next to no engineering by hand [23]. Unsupervised learning has had a catalytic impact on the growing interest in deep learning, but the contributions of solely supervised learning have since been overshadowed. In the longer term, we expect unsupervised learning to become even more significant.

3. Evaluation

We have discussed several supervised and unsupervised text classification methods based on machine and deep learning models so far; however, the shortage of uniform data collection procedures is a big issue when testing text classification techniques. Even if there is a standard collection method available, it can generate differences in model results by simply selecting different training and test sets [146]. Moreover, to compare various performance measures used during separate tests, there may be another difficulty related to process evaluation. In general, performance metrics assess attributes of the performance of the classification task and therefore do not necessarily present similar information. Although the underlying mechanics of various measurement metrics vary, it is important to consider precisely what each of these metrics describes and what kind of data they are attempting to express for comparability. Some examples of such performance measures include precision, recall, accuracy, microaverage, macroaverage, and F-measure. These calculations are based on a "confusion matrix" composed of true positive, false positive, false negative, and true negative [47]. Accuracy is considered the fraction of accurate predictions in overall predictions. The fraction of known positives accurately estimated is referred to as recall. The fraction of positives accurately estimated for all positives is called precision.

Another technique for evaluating how well our machine learning models perform on unknown data is cross-validation. If we expose the model to entirely new, previously unknown data, it may not be as accurate in predicting and may fail to generalize over the new data. Overfitting is the term for this issue. Because it is unable to discover patterns, the model does not always train effectively on the training set. It would not do well on the test set in this situation. Underfitting is the term for this issue. We employ cross-validation to solve overfitting issues. A cross-validation is a resampling approach in which the data set is split into two parts: training data and test data. The model is trained using training data, and the model is predicted using test data that has yet to be observed. If the model performs well on the test data and has a high level of accuracy, it has not overfitted the training data and may be used to forecast. K-fold cross-validation is the most basic type of cross-validation. Other types of cross-validation include variations on k-fold cross-validation or entail repeating k-fold cross-validation rounds. The data is initially partitioned into k equally sized segments or folds in k-fold cross-validation. Following that, k iterations of training and validation are undertaken, with each iteration holding out a different fold of the data for validation and the remaining k-1 folds being employed for learning [147]. For each of the k "folds", the following approach is used:

(i) The folds are used as training data to train a model

(ii) The generated model is validated using the remaining data (i.e., it is used as a test set to compute a performance measure such as accuracy)

In cases where there is uncertainty, entropy is especially appealing as a predictor of classification quality. It denotes how well the class membership probabilities are distributed throughout the specified classes. Its usefulness as a predictor of classification accuracy is predicated on the notion that in an accurate classification, each sentence has a high likelihood of belonging to just one class [159]. Cross-entropy loss is a key indicator for evaluating the performance of a classification issue. The prediction is a probability vector, which means that it reflects the anticipated probabilities of all classes, which add up to one. In a neural network, this is commonly accomplished by activating the final layer with a softmax function, but anything goes, it just has to be a probability vector. Maximum entropy is used in our text classification scenario to estimate the conditional distribution of a class label given a document. A collection of word-count characteristics represents a document. On a class-by-class basis, the labeled training data is utilized to estimate the anticipated value of these word counts. Iterative scaling is improved to obtain a text classifier with an exponential shape that is compatible with the limitations of the labeled data [160].

4. Comparative Analysis

The following section summarizes the benefits and limitations of feature extraction, feature selection methods, and supervised and unsupervised machine and deep learning models used for a text classification task.

TABLE 1: Benefits and limitations of text representation or feature extraction methods.

Method	Benefits	Limitations
Bag-of-words	Works well with unseen words and is easy to implement as it is based on the most frequent terms in a document	Does not cover the syntactic and semantic relation of the words, common words impact classification
TF-IDF	In like bag-of-words approach, common words are excluded due to IDF so does not impact the result	Does not cover the syntactic and semantic relation of the words
Word2vec	Covers the syntactic and semantic relation of the words in the text	Does not cover the words' polysemy
GloVe	As the same as word2vec but performs better, eliminates common words, trained on a large corpus	Does not cover the words' polysemy and does not work well for unseen words
Context-aware representation	Covers the context or meaning of the words in the text	Huge memory is required for storage and does not work well for unseen words

TABLE 2: Benefits and limitations of feature selection methods.

Method		Benefits	Limitations
Univariate filter method	Information gain	Results into the relevance of an attribute or feature	Biased towards multi-valued attributes and overfitting
	Chi-square	Reduces training time and avoids overfitting	Highly sensitive to sample size
	Fishers' score	Evaluates features individually to reduce the feature set	Does not handle features redundancy
	Pearson's correlation coefficient	Is simplest and fast and measures the linear correlation between features	It is only sensitive to a linear relationship
	Variance threshold	Removes features with variance below a certain cutoff	Does not consider the relationship with the target variable
Multi-variate filter method	mRMR (minimal redundancy maximum relevance)	Measures the nonlinear relationship between feature and target variable and provides low error accuracies	Features may be mutually as dissimilar to each other as possible
	Multi-variate relative discriminative criterion	Best determines the contribution of individual features to the underlying dimensions	Does not fit for a small sample size
Linear multi-variate wrapper method	Recursive feature elimination	Considers high-quality top-N features and removes weakest features	Computationally expensive and correlation of features not considered
	Forward/backward stepwise selection	Is computationally efficient and greedy optimization	Sometimes impossible to find features with no correlation between them
	Genetic algorithm	Accommodates data set with a large number of features and knowledge about a problem not required	Stochastic nature and computationally expensive
Nonlinear multi-variate wrapper methods	Nonlinear kernel multiplicative	De-emphasizes the least useful features by multiplying features with a scaling factor	The complexity of kernel computation and multiplication
	Relief	Is feasible for binary classification, based on nearest neighbor instance pairs and is noise-tolerant	Does not evaluate boundaries between redundant features, not suitable for the low number of training data sets
Embedded methods	LASSO	L1 regularization reduces overfitting, and it can be applied when features are even more than the data	Random selection when features are highly correlated
	Ridge regression	L2 regularization is preferred over L1 when features are highly correlated	Reduction of features is a challenge
	Elastic net	Is better than L1 and L2 for dealing with highly correlated features, is flexible, and solves optimization problems	High computational cost

4.1. Comparative Analysis of Standard Text Representation or Feature Extraction Methods. The two major feature extraction methods are highlighted: weighted words and embedding of words. By considering their frequency and co-occurrence details, word embedding approaches learn through sequences of words. These strategies are also unsupervised models to create word vectors. In comparison, the properties of weighted terms are based on counting words in documents and can be used as a basic word representation ranking scheme. Each approach poses specific constraints. Weighted words explicitly quantify text similarities from the word-count space, which enhances the computational time for large vocabulary. Although counting unique terms offers

TABLE 3: Benefits and limitations of the machine and deep learning model.

Model	Benefits	Limitations
Naïve Bayes	It needs less training data; probabilistic approach handles continuous and discrete data; and it is not sensitive to irrelevant features, easily updatable	Data scarcity can lead to loss of accuracy because it is based on assumption that any two features are independent given the output class.
SVM	It is possible to apply to unstructured data also such as text, images, and so on; kernel provides strength to the algorithm and can work for high-dimensional data	It needs long training time on large data sets and is difficult to choose good kernel function, and choosing key parameters varies from problem to problem.
KNN	It can be implemented for classification and regression problems and produces the best results if large training data is available or even noisy training data, preferred for multi-class problems	Cost is high for computing distance for each instance; finding attributes for distance-based learning is quite a difficult task; imbalanced data causes problems; and no treatment is required for missing value.
Decision tree	It reduces ambiguity in decision-making; implicitly performs feature selection, easy representation, and interpretation; and requires fewer efforts for data preparation	It is unstable due to the effect of changes in data requires changes in the whole structure, is not suitable for continuous values, and causes overfitting problem.
Boosted decision tree	It is highly interpretable and prediction accuracy is improved. It can model feature interactions and execute feature selection on its own. Gradient boosted trees are less prone to overfitting since they are trained on a randomly selected subset of the training data.	These are computationally expensive and frequently need a large number of trees (>1,000), which can take a long time and consume a lot of memory.
Random forest	In contrast to other methods, clusters of decision trees are very easy to train, and the preparation and preprocessing of the input data do not require.	More trees in random forests increase the time complexity in the prediction stage, and high chances of overfitting occur.
CNN	It provides fast predictions, is best suited for a large volume of data, and requires no human efforts for feature design.	Computationally expensive requires a large data set for training.
RNN	It implements feedback model so considers best for time series problems and makes accurate predictions than other ANN models.	Training of model is difficult and takes a long time to find nonlinearity in data, and gradient vanishing problem occurs.
LSTM, Bi-LSTM	Adds short- and long-term memory components into RNN so it considers best for applications that have a sequence and uses for solving NLP problems such as text classification and text generation, and computing speed is high. Bi-LSTM solves the issue of predicting fixed sequence to sequence.	It is expensive and complex due to the backpropagation model, increases the dimensionality of the problem, and makes it harder to find the optimal solution. Since Bi-LSTM has double LSTM cells so it is expensive to implement.
Gated RNN (GRU)	In natural language processing, GRUs learn quicker and perform better than LSTMs on less training data. As it requires fewer training parameters. GRUs are simpler and hence easier to modify and do not need memory units, such as by adding extra gates if the network requires more input.	Slow convergence and limited learning efficiency are still issues with GRU.
Transformer with an attention mechanism	The issue with RNNs and CNNs is that when sentences are too long, they are not able to keep up with context and content. By paying attention to the word that is currently being operated on this limitation was resolved, the attention strategy is an effort to selectively concentrate on a few important items while avoiding those in deep neural networks to execute the same operation, enabling much more parallelization than RNNs and thus reduces training times.	At inference time, it is strongly compute-intensive.

independent confirmation of similarity, semantic comparisons between words are not taken into consideration [148]. Word embedding techniques solve this challenge but are constrained by the need for a large corpus of text data sets to train [21]. Therefore, researchers tend to use vectors with pretrained word embedding [149]. Table 1 presents the advantages and limitations of each technique of text representation or feature extraction.

4.2. Comparative Analysis of Standard Feature Selection Methods. Some studies have preferred Fisher's linear discriminant for the support-vector-based feature ranking model [99, 100]. The author has mentioned that the filter-based method provides distinguish feature selector that has further selected distinctive features that possess term characteristics during the elimination of uninformative ones [92]. It offered performance by decreasing processing time and

increasing classification accuracy. Another technique in the wrapper method has gained popularity fast forward selection technique for selecting the best subset of a feature that demanded less computational effort as compared to other methods [107]. In the genetic algorithmic approach, a chaos genetic algorithm was proposed to simplify the feature selection method and obtained higher accuracy of classification technique [108, 109, 161]. Embedded methods were preferred over wrappers in many studies. Researchers mentioned that embedded methods have performed better than wrapper computationally, but these algorithms perform selection features as a subpart of the learning technique. Furthermore, these were followed by hybrid feature selection approaches where both filter and wrapper methods were combined, and these approaches were proven more computationally effective than the performance of a single selection technique. It was observed by the researchers that sometimes original features may not be optimal dimensions for text document representation. They provided text feature extraction methods to solve the problem by creating new feature space or artificial terms [117]. Table 2 presents the benefits and limitations of feature selection methods.

4.3. Comparison of State-of-the-Art Machine and Deep Learning Models for Text Classification. The performance of the classifier depends on the selection of feature selection and extraction method. The supervised and unsupervised machine learning techniques have offered a variety of classifiers that performed well in a variety of domain-specific classification problems. The following Table 3 presents their pros and cons. Recent studies have focused on deep learning or neural network-based classifiers such as CNN, RNN, and RNN with LSTM that have shown better results than conventional algorithms such as SVM and KNN in solving a different range of problems. RNN and LSTM have been used in many NLP applications due to their capacity of capturing complex patterns within the text [151]. It has also been used in sequence labeling tasks in POS (part of speech) activity. But it has offered a lot of future scope in resolving complexities involved in the backpropagation technique used in RNN and making the learning model cheaper and faster [104].

5. Research Gap and Future Direction

From this review of existing studies, we identified that there exist certain gaps, which we can plan to fill in the future. While labeling unstructured text data manually, it takes a lot of time to understand the data to categorize it. Also, it needs specialists for understanding domain-specific data. In such scenarios, the machine learning algorithms do not produce the expected accuracy [12]. Extraction of meaning or finding semantic relations between words of unstructured data is a complex task, which has been tried by several authors using NLP techniques for years [24, 162]. However, these methods prove inefficient when pursuing building a high-quality classification system. It offers opportunities to design semisupervised machine learning models to label some parts

of training data manually, and the rest data can be trained using machine learning algorithms [163].

Apart from conventional methods used for representing data sets for extracting patterns from the text data using vector representation based on word-embedding and paragraph-embedding [38, 58], the authors have presented deep neural network models for NLP-based applications using character-embedding [164], unsupervised techniques based on transfer learning, Bert, with fine-tuning with domain-specific data [156]. However, these representations and globally available representations cannot be generalized to unseen texts, which are very specific to a particular domain [59]. It offers opportunities to design methodology for extending the vocabulary of existing representations for specific domains.

The deep learning algorithms are proven good in decision-making for NLP-based applications, and these models cannot handle symbols directly [165]. Also, the computational cost of training such algorithms is very high. It offers scope for designing deep neural network-based architecture, which can be inputted with linguistic knowledge, lexical knowledge, and word knowledge from different domains.

6. Conclusions

In this paper, we have provided a detailed review of the complete process of the text classification system. This paper covered various algorithms or methods used in subtasks of classification. It has presented the techniques for data collection from several online sources. The documents were represented with basic techniques and followed by recent research in document representation for different areas of machine learning and natural language processing. To provide suitable and fast classifiers the higher dimensional space of data was reduced to a lower space using feature selection and feature extraction methods. Different algorithms perform differently depending on the domain-specific data collections and to train machine learning text classifiers. The authors have used these algorithms based on the problem statement, and none of the algorithms has proven perfect for all types of problems and data dimensionality.

It is observed that in recent years, some studies focused on new applications of text classification such as multi-label classification [166, 167], and hierarchical classification [168, 169] in the field of natural language processing or machine translation and medical sciences, respectively. The neural network-based algorithms are commonly used in NLP-based problems [136, 137, 142]. However, these algorithms mainly focus on generic data. Moreover, CNN is preferably used for image processing, and RNN with LSTM is preferred for time-series problems. Initially, these algorithms were not applied to text data, but in recent years, CNN with character-embedding is being used for document representation and feature selection in text documents [58]. The transformers-based unsupervised models also overcome the issue of RNN and LSTM, but these techniques are proven computationally expensive. With the advancement of deep neural networks, we may find in the future that these deep

neural networks will be applied efficiently in the automatic monitoring of web-based text data and classifying unseen data into automated labels [7].

Data Availability

The data used in this study are publicly available.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] A. Atkins, M. Niranjan, and E. Gerding, "Financial news predicts stock market volatility better than close price," *The Journal of Finance and Data Science*, vol. 4, no. 2, pp. 120–137, 2018.

[2] J. Robinson, A. Glean, and W. Moore, "How does news impact on the stock prices of green firms in emerging markets?" *Research in International Business and Finance*, vol. 45, pp. 446–453, 2018.

[3] G. Löffler, L. Norden, and A. Rieber, *Salient News and the Stock Market Impact of Tone in Rating Reports*, Ssrn, Rochester, NY, USA, pp. 1–41, 2016.

[4] M. N. Elagamy, C. Stanier, and B. Sharp, "Stock market random forest-text mining system mining critical indicators of stock market movements," in *Proceedings of the 2018 Second International Conference on Natural Language and Speech Processing (ICNLSP)*, pp. 1–8, IEEE, Algiers Algeria, June 2018.

[5] Y. Zhang, Y. Dang, H. Chen, M. Thurmond, and C. Larson, "Automatic online news monitoring and classification for syndromic surveillance," *Decision Support Systems*, vol. 47, no. 4, pp. 508–517, 2009.

[6] J. C. West, D. E. Clarke, F. F. Duffy et al., "Availability of mental health services prior to health care reform insurance expansions," *Psychiatric Services*, vol. 67, no. 9, pp. 983–989, 2016.

[7] M. Ali, S. Khalid, M. I. Rana, and F. Azhar, "A probabilistic framework for short text classification," in *Proceedings of the IEEE Eighth Annu Comput Commun Work Conf CCWC 2018*, pp. 742–747, Las Vegas, NV, USA, February 2018.

[8] C. Romero and S. Ventura, "Educational data mining: a survey from 1995 to 2005," *Expert Systems with Applications*, vol. 33, no. 1, pp. 135–146, 2007.

[9] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6527–6535, 2009.

[10] D. Thorleuchter and D. Van Den Poel, "Predicting e-commerce company success by mining the text of its publicly-accessible website," *Expert Systems with Applications*, vol. 39, no. 17, Article ID 13026, 2012.

[11] G. Yang, M. A. Jan, A. U. Rehman, M. Babar, M. M. Aimal, and S. Verma, "Interoperability and data storage in Internet of multimedia things: investigating current trends, research challenges and future directions," *IEEE Access*, vol. 8, Article ID 124382, 2020.

[12] A. Gandomi and M. Haider, "Beyond the hype: big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144, 2015.

[13] Y. Sun, A. K. C. Wong, and M. S. Kamel, "Classification of imbalanced data: a review," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 4, pp. 687–719, Lyon, France, 2009.

[14] S. Manne and S. S. Fatima, "A novel approach for text categorization of unorganized data based with information extraction," *International Journal of Computational Science and Engineering*, vol. 3, pp. 2846–2854, 2011.

[15] B. S. Harish, D. S. Guru, and S. Manjunath, "Representation and classification of text documents: a brief review," *IJCA, Spec Issue Recent Trends Image Process Pattern Recognit*, pp. 110–119, 2010.

[16] Y. Liang, Y. Liu, C. K. Kwong, and W. B. Lee, "Learning the "Whys": discovering design rationale using text mining - an algorithm perspective," *Computer-Aided Design*, vol. 44, no. 10, pp. 916–930, 2012.

[17] B. Liu, X. Li, W. S. Lee, and P. S. Yu, "Text classification by labeling words," *Artificial Intelligence*, vol. 34, pp. 425–430, 2004.

[18] D. Y. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," *Advances in Neural Information Processing Systems*, vol. 16, pp. 321–328, 2004.

[19] H. U. Rehman, R. Rafique, and M. C. M. Nasir, "Forecasting CO2 emissions from energy, manufacturing and transport sectors in Pakistan: statistical vs. Machine learning methods," *Mach Learn Methods*, vol. 28, p. 21, 2017.

[20] N. Kundu, G. Rani, V. S. Dhaka et al., "IoT and interpretable machine learning based framework for disease prediction in pearl millet," *Sensors*, vol. 21, no. 16, p. 5386, 2021.

[21] Y. Bengio, R. Ducharme, and P. Vincent, "A neural probabilistic language model," *Advances in Neural Information Processing Systems*, vol. 3, pp. 1–11, 2001.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[23] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[24] A. Ittoo, L. M. Nguyen, and A. Van Den Bosch, "Text analytics in industry: challenges, desiderata and trends," *Computers in Industry*, vol. 78, pp. 96–107, 2016.

[25] F. Da Costa Albuquerque, M. A. Casanova, J. A. F. De Macedo, M. T. M. de Carvalho, and C. Renso, "A proactive application to monitor truck fleets," in *Proceedings of the 2013 IEEE Fourteenth International Conference on Mobile Data Management*, vol. 1, pp. 301–304, IEEE, Milan, Italy, July 2013.

[26] S. Schmidt, S. Schnitzer, and C. Rensing, "Text classification based filters for a domain-specific search engine," *Computers in Industry*, vol. 78, pp. 70–79, 2016.

[27] A. Hussain, S. Nazir, F. Khan et al., "A resource efficient hybrid proxy mobile IPv6 extension for next generation IoT networks," *IEEE Internet of Things Journal*, p. 1, 2021.

[28] M. Kumar, P. Mukherjee, K. Verma, S. Verma, and D. B. Rawat, "Improved deep convolutional neural network based malicious node detection and energy-efficient data transmission in wireless sensor networks," *IEEE Transactions on Network Science and Engineering*, p. 1, 2021.

[29] P. Rani, V. S. Kavita, S. Verma, and G. N. Nguyen, "Mitigation of black hole and gray hole attack using swarm inspired algorithm with artificial neural network," *IEEE Access*, vol. 8, Article ID 121755, 2020.

[30] B. Billal, A. Fonseca, and F. Sadat, "Named entity recognition and hashtag decomposition to improve the classification of tweets," in *Proceedings of the Second Workshop on Noisy User generated Text*, pp. 64–73, Osaka, Japan, December 2016.

[31] R. P. Schumaker and H. Chen, "A quantitative stock prediction system based on financial news," *Information Processing & Management*, vol. 45, no. 5, pp. 571–583, 2009.

[32] Y. Chen, Z. Li, L. Nie, X. Hu, and X. Wang, "Supervised bayesian network model for microblog topic classification," vol. 1, pp. 561–576, 2012.

[33] R. Xia, C. Zong, X. Hu, and E. Cambria, "Feature ensemble plus sample selection: domain adaptation for sentiment classification," in *Proceedings of the IJCAI Int Jt Conf Artif Intell 2015*, pp. 4229–4233, AAAI Press, California, CA, USA, January 2015.

[34] N. Majumder, S. Poria, A. Gelbukh, E. Cambria, and E. Cambria, "Deep learning-based document modeling for personality detection from text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, 2017.

[35] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.

[36] L. Gaur, G. Singh, A. Solanki et al., "Disposition of Youth in Predicting Sustainable Development Goals Using the Neuro-Fuz," *Human-Centric Computing and Information Sciences*, vol. 11, pp. 2192–1962, 2021.

[37] R. H. W. Pinheiro, G. D. C. Cavalcanti, and I. R. Tsang, "Combining dissimilarity spaces for text categorization," *Information Sciences*, vol. 406-407, pp. 87–101, 2017.

[38] T. Sabbah, A. Selamat, M. H. Selamat et al., "Modified frequency-based term weighting schemes for text classification," *Applied Soft Computing*, vol. 58, pp. 193–206, 2017.

[39] V. Dogra, S. Verma, K. Verma, N. Ghosh, U. Le, and D.-N. au, "A comparative analysis of machine learning models for banking news extraction by multiclass classification with imbalanced datasets of financial news: challenges and solutions," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 7, no. 3, p. 35, 2022.

[40] Y. Zhao, B. Qin, and T. Liu, "Creating a fine-grained corpus for Chinese sentiment analysis," *IEEE Intelligent Systems*, vol. 30, no. 1, pp. 36–43, 2015.

[41] Y. Zhang, R. Jin, and Z. H. Zhou, "Understanding bag-of-words model: a statistical framework," *International Journal of Machine Learning and Cybernetics*, vol. 1, no. 1-4, pp. 43–52, 2010.

[42] T. Joseph and Y. B. Lev Ratinov, "Word representations: a simple and general method for semi-supervised learning," in *Proceedings of the Forty Eighth Annu Meet Assoc Comput Linguist Assoc Comput Linguist*, pp. 384–394, AAAI Press, California, CA, USA, July 2010.

[43] R. Silipo and K. Melcher, *Text Encoding: A Review*, Kdnuggets, 2019, https://www.kdnuggets.com/2019/11/text-encoding-review.html.

[44] Z. Sivic and Zisserman, "Video Google: a text retrieval approach to object matching in videos," in *Proceedings of the Ninth IEEE International Conference on Computer Vision*, vol. 2, pp. 1470–1477, IEEE, New York, NY, USA, October 2003.

[45] Li-P. Jing, H.-K. Huang, and H.-Bo Shi, "Improved Feature Selection Approach TFIDF in Text Mining," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, vol. 2, pp. 944–946, IEEE, New York, NY, USA, November 2002.

[46] E. Montañés, I. Díaz, J. Ranilla, E. Combarro, and J. Fernandez, "Scoring and selecting terms for text categorization," *IEEE Intelligent Systems*, vol. 20, no. 3, pp. 40–47, 2005.

[47] Z. Li, S. Verma, and M. Jin, "Power allocation in massive MIMO-HWSN based on the water-filling algorithm," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8719066, 11 pages, 2021.

[48] Y. Bengio, R. Ducharme, V. Pascal, and J. Christen, "A neural probabilistic language model," *Journal of Machine Learning Research*, vol. 3, pp. 1137–1155, 2003.

[49] Y. Xiong, S. Chen, H. Qin et al., "Distributed representation and one-hot representation fusion with gated network for clinical semantic textual similarity," *BMC Medical Informatics and Decision Making*, vol. 20, no. S1, pp. 72–77, 2020.

[50] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: a survey," *Information*, vol. 10, no. 4, pp. 150–168, 2019.

[51] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *Proceedings of the First Int Conf Learn Represent ICLR 2013 - Work Track Proc 1–12*, Scottsdale, Arizona, May 2013.

[52] B. Liu, "Text sentiment analysis based on CBOW model and deep learning in big data environment," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 2, pp. 451–458, 2020.

[53] H. Schwenk, "Continuous space language models," *Computer Speech & Language*, vol. 21, no. 3, pp. 492–518, 2007.

[54] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds., pp. 3111–3119, Curran Associates Inc, Red Hook, NY, USA, 2013.

[55] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 135–146, 2017.

[56] J. Pennington and C. D. M. Richard Socher, "GloVe: global vectors for word representation," in *Proceedings of the 2014 Conf Empir Methods Nat Lang Process (EMNLP)*, pp. 1532–1543, Doha, Qatar, October 2014.

[57] C.-P. Wei and Y.-H. Lee, "Event detection from online news documents for supporting environmental scanning," *Decision Support Systems*, vol. 36, no. 4, pp. 385–401, 2004.

[58] Q. Zhu, X. Li, A. Conesa, and C. Pereira, "GRAM-CNN: a deep learning approach with local context for named entity recognition in biomedical text," *Bioinformatics*, vol. 34, no. 9, pp. 1547–1554, 2018.

[59] V. Prokhorov, M. T. Pilehvar, P. Lio, and N. Collier, "Unseen word representation by aligning heterogeneous lexical semantic spaces," in *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, Honolulu, HI, USA, February 2019.

[60] A. Makazhanov and D. Rafiei, "Predicting political preference of Twitter users," in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 298–305, Association for Computing Machinery, New York NY USA, August 2013.

[61] C. Bosco, V. Patti, and A. Bolioli, "Developing corpora for sentiment analysis: the case of irony and senti-TUT," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 55–63, 2013.

[62] M. E. Peters, M. Neumann, M. Iyyer et al., "Deep contextualized word representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 2227–2237, New Orleans, LA, USA, June 2018.

[63] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep learning--based text classification," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–40, 2022.

[64] X. Zhu, P. Sobihani, and H. Guo, "Long short-term memory over recursive structures," in *Proceedings of the 32nd International Conference on International Conference on Machine Learning*, pp. 1604–1612, Lille, France, July 2015.

[65] P. Zhou, Z. Qi, S. Zheng, J. Xu, and H. Bao, "Text classification improved by integrating bidirectional lstm with two-dimensional max pooling," 2016, https://arxiv.org/abs/1611.06639.

[66] W. Wei, J. Li, L. Cao, Y. Ou, and J. Chen, "Effective detection of sophisticated online banking fraud on extremely imbalanced data," *World Wide Web*, vol. 16, no. 4, pp. 449–475, 2013.

[67] S. I. Alfuraih, N. T. Sui, and D. McLeod, "Using Trusted Email to Prevent Credit Card Frauds in Multimedia Products," *World Wide Web*, vol. 5, pp. 245–256, 2004.

[68] D. Yan and S. Guo, "Leveraging contextual sentences for text classification by using a neural attention model," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 8320316, 11 pages, 2019.

[69] Z. Dai, Z. Yang, Y. Yang, C. Jaime, V. L. Quoc, and S. Ruslan, "Transformer-XL: Attentive language models beyond a fixed-length context," in *Proceedings of the ACL 2019 - Fifty Seventh Annu Meet Assoc Comput Linguist Proc Conf*, pp. 2978–2988arXiv:1901.02860, Florence, Italy, July 2020.

[70] Y. Hu, J. Ding, Z. Dou, and H. Chang, "Short-text classification detector: a bert-based mental approach," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8660828, 11 pages, 2022.

[71] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," 2018, https://arxiv.org/abs/1810.04805.

[72] J. Howard and S. Ruder, "Universal Language Model Fine-tuning for Text Classification," 2018, https://arxiv.org/abs/1801.06146.

[73] Y. Sun, S. Wang, Y. Li et al., "Enhanced representation through knowledge integration," 2019, https://arxiv.org/abs/1904.09223.

[74] M. T. Varun Dogra Assvknj and, "Analyzing DistilBERT for sentiment classification of banking financial news," in *Intelligent Computing and Innovation on Data Science*, S.-L. Peng, S.-Y. Hsieh, S. Gopalakrishnan, and Balaganesh, Eds., p. 582, Springer Singapore, Singapore, 2021.

[75] V. Dogra, S. Verma, and A. Singh, "Banking news-events representation and classification with a novel hybrid model using DistilBERT and rule-based features," *Computer Science*, vol. 12, pp. 3039–3054, 2021.

[76] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter," pp. 2–6, 2019, https://arxiv.org/abs/1910.01108.

[77] Z. Ye, Q. Guo, Q. Gan, X. Qiu, and Z. Zhang, "BP-transformer: modelling long-range context via binary partitioning," 2019, https://arxiv.org/abs/1911.04070.

[78] I. Yamada and H. Shindo, "Aip R Neural Attentive Bag-Of-Entities Model for Text Classification," arXiv:1909.01259, 2019.

[79] T. B. Brown, B. Mann, N. Ryder et al., "Language models are few-shot learners," in *Proceedings of the Adv Neural Inf Process Syst 2020*, December 2020.

[80] B. Magnini, C. Strapparava, G. Pezzulo, and A. Gliozzo, "The role of domain information in Word Sense Disambiguation," *Natural Language Engineering*, vol. 8, no. 4, pp. 359–373, 2002.

[81] M. Bilenko, R. Mooney, W. Cohen, P. Ravikumar, and S. Fienberg, "Adaptive name matching in information integration," *IEEE Intelligent Systems*, vol. 18, no. 5, pp. 16–23, 2003.

[82] S. J. Huang, W. Gao, and Z. H. Zhou, "Fast multi-instance multi-label learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 11, pp. 2614–2627, 2019.

[83] M.-L. Zhang and Z.-H. Zhou, "ML-KNN: a lazy learning approach to multi-label learning," *Pattern Recognition*, vol. 40, no. 7, pp. 2038–2048, 2007.

[84] J. Jun Yan, B. Benyu Zhang, N. Ning Liu et al., "Effective and efficient dimensionality reduction for large-scale and streaming data preprocessing," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 3, pp. 320–333, 2006.

[85] X. Xu, T. Liang, J. Zhu, D. Sun, and T. au, "Review of classical dimensionality reduction and sample selection methods for large-scale data processing," *Neurocomputing*, vol. 328, pp. 5–15, 2019.

[86] N. Armanfard, J. P. Reilly, and M. Komeili, "Local feature selection for data classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 6, pp. 1217–1227, 2016.

[87] S. Pölsterl, S. Conjeti, N. Navab, and A. Katouzian, "Survival analysis for high-dimensional, heterogeneous medical data: exploring feature extraction as an alternative to feature selection," *Artificial Intelligence in Medicine*, vol. 72, pp. 1–11, 2016.

[88] D. Mladenić and M. Grobelnik, "Feature selection on hierarchy of web documents," *Decision Support Systems*, vol. 35, no. 1, pp. 45–87, 2003.

[89] A.-C. Haury, P. Gestraud, and J.-P. Vert, "The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures," *PLoS One*, vol. 6, no. 12, Article ID e28210, 2011.

[90] J. Yang, Y. Liu, X. Zhu, Z. Zhang, Z. Liu, and X. Zhang, "A new feature selection based on comprehensive measurement both in inter-category and intra-category for text categorization," *Information Processing & Management*, vol. 48, no. 4, pp. 741–754, 2012.

[91] I. Inza, P. Larrañaga, R. Blanco, and A. J. Cerrolaza, "Filter versus wrapper gene selection approaches in DNA microarray domains," *Artificial Intelligence in Medicine*, vol. 31, no. 2, pp. 91–103, 2004.

[92] A. K. Uysal and S. Gunal, "A novel probabilistic feature selection method for text classification," *Knowledge-Based Systems*, vol. 36, pp. 226–235, 2012.

[93] D. Roobaert, G. Karakoulas, and N. V. Chawla, "Information gain, correlation and support vector machines," *Feature Extraction*, vol. 470, pp. 463–470, 2008.

[94] S. Lei, "A feature selection method based on information gain and genetic algorithm," in *Proceedings of the 2012 International Conference on Computer Science and Electronics*

Engineering, vol. 2, pp. 355–358, Hangzhou, China, March 2012.

[95] X. Jin, A. Xu, R. Bie, and P. Guo, "Machine learning techniques and chi-square feature selection for cancer classification using SAGE gene expression profiles," *Lecture Notes in Computer Science*, Springer, Berlin, Germany, pp. 106–115, 2006.

[96] Y.-T. Chen and M. C. Chen, "Using chi-square statistics to measure similarities for text categorization," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3085–3090, 2011.

[97] Q. Gu, Z. Li, and J. Han, "Generalized Fisher score for feature selection," *A brief review of Fisher score. Ratio*, AUAI Press, Arlington, VA, USA, 2010.

[98] J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio, and V. Vapnik, "Feature selection for SVMs," *Advances in Neural Information Processing Systems*, vol. 13, pp. 668–674, 2000.

[99] E. Youn, L. Koenig, M. K. Jeong, and S. H. Baek, "Support vector-based feature selection using Fisher's linear discriminant and Support Vector Machine," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6148–6156, 2010.

[100] Y. Wang and X.-J. Wang, "A new approach to feature selection in text classification," in *Proceedings of the 2005 International Conference on Machine Learning and Cybernetics*, vol. 6, pp. 3814–3819, Guangzhou, China, August 2005.

[101] A. Genkin, D. D. Lewis, and D. Madigan, "Large-scale bayesian logistic regression for text categorization," *Technometrics*, vol. 49, no. 3, pp. 291–304, 2007.

[102] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[103] M. Labani, P. Moradi, F. Ahmadizar, and M. Jalili, "A novel multivariate filter method for feature selection in text classification problems," *Engineering Applications of Artificial Intelligence*, vol. 70, pp. 25–37, 2018.

[104] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fMRI data," *Trends in Cognitive Sciences*, vol. 10, no. 9, pp. 424–430, 2006.

[105] L. Wang, Y. Lei, Y. Zeng, L. Tong, and B. Yan, "Principal feature analysis: a novel voxel selection method for fMRI data," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 645921, 7 pages, 2013.

[106] P. M. Granitto, C. Furlanello, F. Biasioli, and F. Gasperi, "Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products," *Chemometrics and Intelligent Laboratory Systems*, vol. 83, no. 2, pp. 83–90, 2006.

[107] K. Z. Mao, "Fast orthogonal forward selection algorithm for feature subset selection," *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1218–1224, 2002.

[108] H. Chen, W. Jiang, C. Li, and R. Li, "A heuristic feature selection approach for text categorization by using chaos optimization and genetic algorithm," *Mathematical Problems in Engineering*, vol. 2013, Article ID 524017, 6 pages, 2013.

[109] R. Leardi, R. Boggia, and M. Terrile, "Genetic algorithms as a strategy for feature selection," *Journal of Chemometrics*, vol. 6, no. 5, pp. 267–281, 1992.

[110] I. Guyon, H.-M. Bitter, Z. Ahmed, M. Brown, and J. Heller, "Multivariate non-linear feature selection with kernel multiplicative updates and gram-schmidt relief," in *Proceedings of the BISC Flint-CIBI 2003 Workshop*, pp. 1–11, Berkeley, CA, 2003.

[111] R. J. Urbanowicz, M. Meeker, W. L Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: introduction and review," *Journal of Biomedical Informatics*, vol. 85, pp. 189–203, 2018.

[112] R. J. Urbanowicz, R. S. Olson, P. Schmitt, M. Meeker, and J. H. Moore, "Benchmarking relief-based feature selection methods for bioinformatics data mining," *Journal of Biomedical Informatics*, vol. 85, pp. 168–188, 2018.

[113] N. Mimouni and T. Y. Yeung, "Comparing Performance of Text Pre-processing Methods for Predicting a Binary Position by LASSO Experiment with Textual Data of European Union Public Consultation," in *Proceedings of the Workshop on Women in Data Science 2018*, pp. 18–21, 2018.

[114] B. J. Marafino, W. J. Boscardin, and R. A Dudley, "Efficient and sparse feature selection for biomedical text classification via the elastic net: application to ICU risk stratification from nursing notes," *Journal of Biomedical Informatics*, vol. 54, pp. 114–120, 2015.

[115] P. Taylor, A. E. Hoerl, and R. W. Kennard, *Technometrics Ridge Regression: Biased Estimation for Nonorthogonal Problems Ridge Regression : Biased Estimation Nonorthogonal Problems*, pp. 37–41, 2012.

[116] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B*, vol. 67, no. 2, pp. 301–320, 2005.

[117] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1–47, 2002.

[118] J. C. Gomez and M. F. Moens, "PCA document reconstruction for email classification," *Computational Statistics & Data Analysis*, vol. 56, no. 3, pp. 741–751, 2012.

[119] D. M. Blei, A. Y. Ng, and M. I. Jordan, *Latent Dirichlet Allocation*, vol. 3, pp. 993–1022, 2003.

[120] C. Ordun, S. Purushotham, and E. Raff, "Exploratory Analysis of Covid-19 Tweets Using Topic Modeling, UMAP, and DiGraphs," 2020, https://arxiv.org/abs/2005.03082.

[121] L. Mcinnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," 2020, https://arxiv.org/abs/1802.03426.

[122] H. U. Rehman, M. Shafiq, S. Baig, and U. Manzoor, "Analyzing the epidemiological outbreak of COVID-19," *A Vis Explor data Anal approach J Med Virol*, vol. 92, no. 6, 2021.

[123] H. Cheng and R. Yu, *Text Classification Model Enhanced by Unlabeled Data for LaTeX Formula*, 2021.

[124] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5432–5435, 2009.

[125] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based framework for text categorization," *Procedia Engineering*, vol. 69, pp. 1356–1364, 2014.

[126] C. Lee and G. G. Lee, "Information gain and divergence-based feature selection for machine learning-based text categorization," *Information Processing & Management*, vol. 42, no. 1, pp. 155–165, 2006.

[127] L. Khreisat, "A machine learning approach for Arabic text classification using N-gram frequency statistics," *Journal of Informetrics*, vol. 3, no. 1, pp. 72–77, 2009.

[128] J. Liu, T. Jin, K. Pan, Y. Yang, Y. Wu, and X. Wang, "An improved KNN text classification algorithm based on Simhash," in *Proceedings of the 2017 IEEE 16th Int Conf Cogn Informatics Cogn Comput ICCI∗CC*, pp. 92–95, Oxford UK, July 2017.

[129] C. Apté and S. Weiss, "Data Mining with Decision Trees and Decision Rules," *Future Generation Computer Systems*, vol. 13, 1997.

[130] V. N. Phu, V. T. N. Tran, V. T. N. Chau, N. D. Duy, and K. L. D. au, "A decision tree using ID3 algorithm for English semantic analysis," *International Journal of Speech Technology*, vol. 20, no. 3, pp. 593–613, 2017.

[131] C. N. Mahender, *TEXT CLASSIFICATION AND CLASSIFIERS*, vol. 3, pp. 85–99, 2012.

[132] Y. Yang, "An evaluation of statistical approaches to text categorization," *Inf Retr Boston*, vol. 1, pp. 69–90, 1997.

[133] B. P. Roe, H. J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, "Boosted decision trees as an alternative to artificial neural networks for particle identification," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 543, no. 2-3, pp. 577–584, 2005.

[134] J. Ye, J. H. Chow, J. Chen, and Z. Zheng, "Stochastic gradient boosted distributed decision trees," in *Proceedings of the Eighteenth ACM conference on Information and knowledge management - CIKM '09*, pp. 2061–2064, Hong Kong China, November 2009.

[135] E. Wiener, J. Pedersen, and A. Weigend, "A neural network approach to topic spotting," in *Proceedings of the Annu Symp Doc Anal Inf Retr*, pp. 317–332, 1995.

[136] R. Johnson and T. Zhang, "Effective Use of Word Order for Text Categorization with Convolutional Neural Networks," pp. 103–112, 2014, https://arxiv.org/abs/1412.1058.

[137] Y. Zhang, Z. Zhang, D. Miao, and J. Wang, "Three-way enhanced convolutional neural networks for sentence-level sentiment classification," *Information Sciences*, vol. 477, pp. 55–64, 2019.

[138] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 655–665, Baltimore, Maryland, June 2014.

[139] J. L. Elman, "Finding structure in time," *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.

[140] V. Makarenkov, I. Guy, N. Hazon, T. Meisels, B. Shapira, and L. Rokach, "Implicit dimension identification in user-generated text with LSTM networks," *Information Processing & Management*, vol. 56, no. 5, pp. 1880–1893, 2019.

[141] L. Sarmento, P. Carvalho, M. J. Silva, and E. de Oliveira, "Automatic creation of a reference corpus for political opinion mining in user-generated content," *Attitude*, vol. 29, 2009.

[142] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A Hybrid Deep Model for Fake News Detection," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, Singapore, November 2017.

[143] S. Dong and C. Liu, "Sentiment classification for financial texts based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9524705, 9 pages, 2021.

[144] Y. Zhu, W. Zheng, and H. Tang, "Interactive dual attention network for text sentiment classification," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 8858717, 11 pages, 2020.

[145] J. Chung, "Gated Recurrent Neural Networks on Sequence Modeling," pp. 1–9, 2014, https://arxiv.org/abs/1412.3555.

[146] D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1422–1432, Lisbon, Portugal, September 2015.

[147] M. Zulqarnain, R. Ghazali, M. G. Ghouse, and M. F. Mushtaq, "Efficient processing of GRU based on word embedding for text classification," *JOIV: International Journal on Informatics Visualization*, vol. 3, no. 4, 2019.

[148] O. Kuchaiev and B. Ginsburg, "Factorization tricks for LSTM networks," in *Proceedings of the Fifth Int Conf Learn Represent ICLR 2017 - Work Track Proc 1–6*, Toulon, France, April 2019.

[149] N. Shazeer, A. Mirhoseini, K. Maziarz et al., "Outrageously large neural networks: the sparsely-gated mixture-of-experts layer," 2017, https://arxiv.org/abs/1701.06538.

[150] B. C. W. Jain, "Attention is not Explanation," 2019, https://arxiv.org/abs/1902.10186.

[151] S. Serrano and N. A. Smith, "Is attention interpretable?" in *Proceedings of the ACL 2019 - 57th Annu Meet Assoc Comput Linguist Proc Conf*, pp. 2931–2951, Florence, Italy, July 2020.

[152] S. Vashishth, S. Upadhyay, G. S. Tomar, and M. Faruqui, "Attention Interpretability Across NLP Tasks," pp. 1–10, 2019, https://arxiv.org/abs/1909.11218.

[153] T. Munkhdalai and H. Yu, "Neural semantic encoders," in *Proceedings of the 15th Conference of the European Chapter of the*, Valencia, Spain, April 2017.

[154] T. Commissariat, "Ask me anything," *Physics World*, vol. 33, no. 3, pp. 53–58, 2020.

[155] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," in *Proceedings of the Adv Neural Inf Process Syst 2017*, pp. 5999–6009, Long Beach, CA, USA, December 2017.

[156] C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to fine-tune BERT for text classification?" *Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*, Springer, Berlin, Germany, pp. 194–206, 2019.

[157] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations," pp. 1–17, 2019, https://arxiv.org/abs/1909.11942.

[158] M. Joshi, D. Chen, Y. Liu, D. S. Weld, L. Zettlemoyer, and O. Levy, "SpanBERT: Improving Pre-training by Representing and Predicting Spans," 2019, https://arxiv.org/abs/1907.10529.

[159] C. Holzhey, F. Larsen, and F. Wilczek, "Geometric and renormalized entropy in conformal field theory," *Nuclear Physics B*, vol. 424, no. 3, pp. 443–467, 1994.

[160] K. Nigam, J. Lafferty, and A. Mccallum, "Using maximum entropy for text classification," *Computet Science*, vol. 80, 1999.

[161] J. Atkinson-Abutridy, C. Mellish, and S. Aitken, "Combining information extraction with genetic algorithms for text mining," *IEEE Intelligent Systems*, vol. 19, no. 3, pp. 22–30, 2004.

[162] A. Bhardwaj, Y. Narayan, Vanraj, Pawan, and M. Dutta, "Sentiment analysis for Indian stock market prediction using sensex and nifty," *Procedia Computer Science*, vol. 70, pp. 85–91, 2015.

[163] D. Kim, D. Seo, S. Cho, and P. Kang, "Multi-co-training for document classification using various document representations: TF-IDF, LDA, and Doc2Vec," *Information Sciences*, vol. 477, pp. 15–29, 2019.

[164] M. H. Steinberg, "Clinical trials in sickle cell disease: adopting the combination chemotherapy paradigm," *American Journal of Hematology*, vol. 83, no. 1, pp. 1–3, 2008.

[165] Z. Xu and J. Sun, "Model-driven deep-learning," *National Science Review*, vol. 5, no. 1, pp. 22–24, 2018.

[166] Y. Chen, B. Xiao, Z. Lin, C. Dai, Z. Li, and L. Yan, "Multi-label text classification with deep neural networks," in *Proceedings of the 2018 Sixth IEEE Int Conf Netw Infrastruct Digit Content*, pp. 409–413, IEEE, Guiyang China, November 2018.

[167] R. B. Pereira, A. Plastino, B. Zadrozny, and L. H. C. Merschmann, "Categorizing feature selection methods for multi-label classification," *Artificial Intelligence Review*, vol. 49, no. 1, pp. 57–78, 2018.

[168] R. A. Stein, P. A. Jaques, and J. F. Valiati, "An analysis of hierarchical text classification using word embeddings," *Information Sciences*, vol. 471, pp. 216–232, 2019.

[169] R. A. Sinoara, C. V. Sundermann, R. M. Marcacini, M. A. Domingues, and S. R. Rezende, "Named Entities as Privileged Information for Hierarchical Text Clustering," in *Proceedings of the Eighteenth International Database Engineering & Applications Symposium, Association for Computing Machinery*, pp. 57–66, New York, NY, USA, July 2014.

Research Article

Statistical Analysis for Competing Risks' Model with Two Dependent Failure Modes from Marshall–Olkin Bivariate Gompertz Distribution

Min Wu [ID],[1] **Fode Zhang** [ID],[2] **Yimin Shi**,[3] **and Yan Wang** [ID][4]

[1]*School of Economics and Management, Shanghai Maritime University, Shanghai 201306, China*
[2]*Center of Statistical Research, School of Statistics, Southwestern University of Finance and Economics, Chengdu 611130, Sichuan, China*
[3]*School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China*
[4]*School of Science, Xi'an Polytechnic University, Xi'an 710072, Shaanxi, China*

Correspondence should be addressed to Min Wu; mwu@shmtu.edu.cn and Fode Zhang; fredzh@swufe.edu.cn

The bivariate or multivariate distribution can be used to account for the dependence structure between different failure modes. This paper considers two dependent competing failure modes from Gompertz distribution, and the dependence structure of these two failure modes is handled by the Marshall–Olkin bivariate distribution. We obtain the maximum likelihood estimates (MLEs) based on classical likelihood theory and the associated bootstrap confidence intervals (CIs). The posterior density function based on the conjugate prior and noninformative (Jeffreys and Reference) priors are studied; we obtain the Bayesian estimates in explicit forms and construct the associated highest posterior density (HPD) CIs. The performance of the proposed methods is assessed by numerical illustration.

1. Introduction

It is extremely common that the failure of a product or a system contains several competing failure modes in reliability engineering; any failure mode will lead to the failure result. Competing risks' data contain the failure time and the corresponding failure mode, which can be modeled by the competing risks' model and has been commonly performed in many research fields, such as engineering and medical statistics. Previous studies have mostly assumed the competing failure modes to be independent; Wang et al. [1], Ren and Gui [2], and Qin and Gui [3] focused on the independent competing risks' model under progressively hybrid censoring from Weibull and Burr-XII distributions. Objective Bayesian analysis for the competing risks' model with Wiener degradation phenomena and catastrophic failures was studied by Guan et al. [4]. In practice, the independency relationship between different failure modes is a very special case; a more common situation is dependency. That is, the failure mechanisms are interactive and interdependent; the occurrence of one failure mode will affect the occurrence of other failure modes. For example, a ship fixed carbon dioxide fire extinguishing system can fail due to pressure gauge, distribution value, cylinder group, and so on; these failure modes are dependent because they all are related to the storage environment. Therefore, it is more reasonable to assume dependency among different competing failure modes. The competing risks' model considers the product or system with multiple dependent competing failure modes, any one of which will cause the occurrence of failure. The dependent competing risks' model has been extensively studied. Zhang et al. [5] and Zhang et al. [6] studied the dependent competing risks' model under accelerated life testing (ALT) by copula function to measure the dependence between different competing failure modes; the results indicate the copula construction method has good accuracy

and universality. Wang and Yan [7] and Wu et al. [8] also studied this model under ALT and progressively hybrid-censoring scheme using Clayton copula and Gumbel copula, respectively. For other related works, see the works of Lo and Wike [9] and Fang et al. [10].

In addition to using copula function to handle the relationship between different competing failure modes, the bivariate or multivariate distribution also can be used to account for the correlation between different failure modes. The Marshall–Olkin distribution [11], which has many good properties, is the best-known bivariate distribution and has been discussed extensively; it has a parameter to describe the dependence structure. Li et al. [12], Kundu and Gupta [13], and Bai et al. [14] provided reviews on Marshall-l–Olkin–Weibull distribution; Kundu and Gupta [13] obtained the explicit forms of the unknown parameters when the shape parameter is known; when the shape parameter is unknown, they used the importance sampling to compute the Bayesian estimates of the unknown parameters. Bai et al. [14] discussed the statistical analysis for the accelerated dependent competing risks' model under Type-II hybrid censoring schemes. Guan et al. [15] studied objective Bayesian analysis for the Marshall–Olkin exponential distribution based on reference priors; they also found that some of the reference priors are also matching priors and the posterior distributions based on these priors are proper.

The Gompertz distribution is a widely used growth model which has been studied extensively; Ismail [16] studied the Bayesian analysis of Gompertz distribution parameters and acceleration factor in the case of partially accelerated life testing under Type-I censoring scheme. Ghitany et al. [17] considered a progressively censored sample from Gompertz distribution; they discussed the existence and uniqueness of the MLEs of the unknown parameters. The Gompertz distribution plays an important role in fitting clinical trials' data in medical science and can be used to the theory of extreme-order statistics. In this paper, we will study the dependent competing risks' model from the Marshall–Olkin bivariate Gompertz (MOGP) distribution, which is a bivariate distribution with Gompertz marginal distributions. We focus our attention on the

statistical analysis of the model parameters, including classical likelihood inference, Bayesian analysis, and objective Bayesian analysis. Because the Bayesian analysis based on conjugate prior is sensitive to the hyperparameters, inappropriate choice of it will cause bad priors. Based on this reason, we propose the objective Bayesian analysis based on noninformative priors for comparison. The objective Bayesian inference has been studied by Guan et al. [14], Bernardo [18], and Berger and Bernardo [19] based on Reference and Jeffreys priors.

In the rest of this paper, we will present the model description and some properties. Section 3 presents the MLEs and associated bootstrap CIs. In Section 4, Bayesian estimates and associated HPD CIs based on conjugate prior, Jeffreys prior [20], and reference priors [18] are obtained, and these priors lead to proper posteriors which are proved. Section 5 presents some results obtained from simulation study and illustrative analysis. Section 6 gives some final concluding remarks.

2. Model Description

Suppose that $f(t; \lambda, \theta)$ is a Gompertz distribution; the density function and reliability function of it are

$$f(t; \lambda, \theta) = \theta e^{(\lambda t - \theta(e^{\lambda t} - 1)/\lambda)}, \ \lambda, \ \theta > 0, \ t > 0,$$
$$S(t; \lambda, \theta) = e^{(-\theta(e^{\lambda t} - 1)/\lambda)}, \ \lambda, \ \theta > 0, \ t > 0, \tag{1}$$

where λ is shape parameter and θ is scale parameter.

Suppose U_0, U_1, and U_2 are three independent Gompertz variables with different scale parameters, that is, $U_0 \sim GP(\lambda, \theta_0)$, $U_1 \sim GP(\lambda, \theta_1)$, and $U_2 \sim GP(\lambda, \theta_2)$. Let $T_1 = \min(U_0, U_1)$ and $T_2 = \min(U_0, U_2)$; we obtain $T_1 \sim GP(\lambda, \theta_0 + \theta_1)$ and $T_2 \sim GP(\lambda, \theta_0 + \theta_2)$. Then, the pair of variables (T_1, T_2) follows the MOGP distribution denoted by $(T_1, T_2) \sim MOGP(\lambda, \theta_0, \theta_1, \theta_2)$. When $\theta_0 = 0$, the two variables T_1 and T_2 are independent and T_1 and T_2 will be dependent when $\theta_0 > 0$; hence, θ_0 can be regarded as a correlation coefficient between T_1 and T_2.

The joint PDF of (T_1, T_2) can be written as

$$f_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2) = \begin{cases} f(t_1; \lambda, \theta_0 + \theta_1) f(t_2; \lambda, \theta_2) & t_1 > t_2 \\ \\ f(t_1; \lambda, \theta_1) f(t_2; \lambda, \theta_0 + \theta_2) & t_1 < t_2. \\ \\ \left(\dfrac{\theta_0}{(\theta_0 + \theta_1 + \theta_2)} \right) f(t; \lambda, \theta_0 + \theta_1 + \theta_2) & t_1 = t_2 = t \end{cases} \tag{2}$$

The surface plots of $f_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)$ are presented in Figure 1. From Figure 1, we can see that $f_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)$ is a unimodal function.

Put n identical products into test, and each product has two dependent failure modes with lifetimes T_1, T_2,

$(T_1, T_2) \sim MOGP(\lambda, \theta_0, \theta_1, \theta_2)$. Then, the system lifetime is $X = \min(T_1, T_2) \sim MOGP(\lambda, \theta_0 + \theta_1 + \theta_2)$. Let $\delta_{0l} = I(T_{1l} = T_{2l})$, $\delta_{1l} = I(T_{1l} < T_{2l})$, and $\delta_{2l} = I(T_{1l} > T_{2l})$, for $l = 1, \cdots, n$, where $I(\cdot)$ is an indicator function. Then, we can compute $n_0 = \sum_l \delta_{0l}$, $n_1 = \sum_l \delta_{1l}$, $n_2 = \sum_l \delta_{2l}$, and $n = n_0 + n_1 + n_2$.
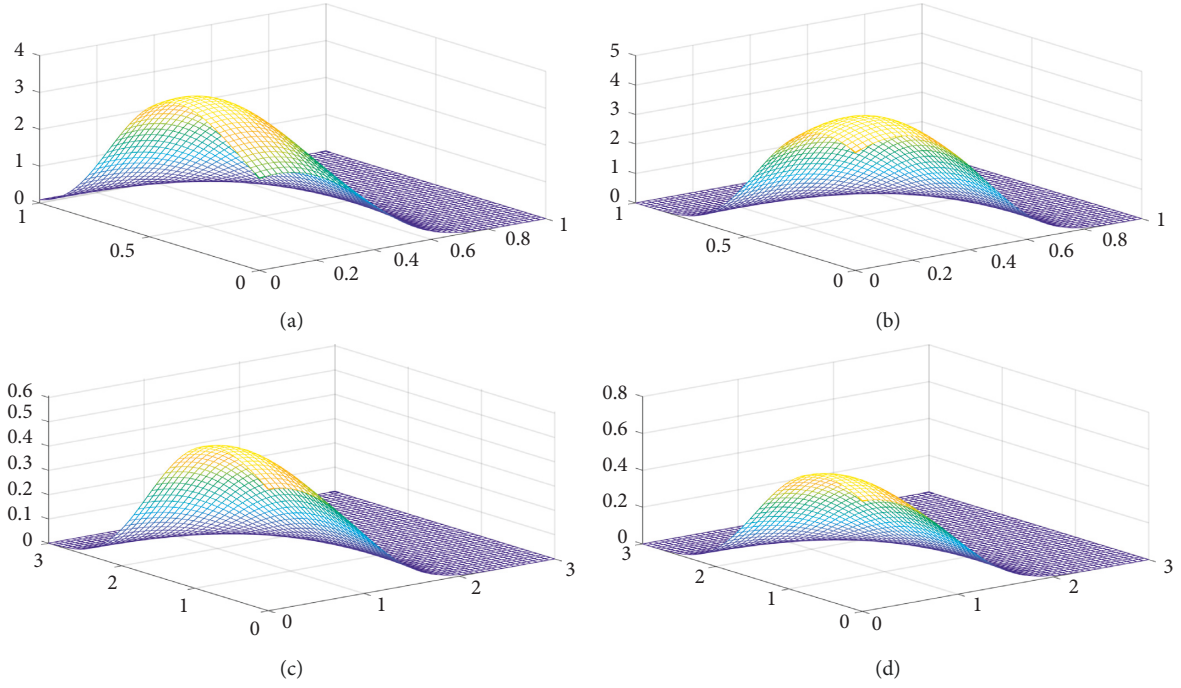
FIGURE 1: Surface plot of $f_{T_1,T_2}(t_1,t_2;\lambda,\theta_0,\theta_1,\theta_2)$ with different values of $\lambda, \theta_0, \theta_1, \theta_2$. (a) $(\lambda, \theta_0, \theta_1, \theta_2) = (3, 0.5, 2, 1)$. (b) $(\lambda, \theta_0, \theta_1, \theta_2) = (3, 1.5, 0.5, 2)$. (c) $(\lambda, \theta_0, \theta_1, \theta_2) = (1, 0.5, 0.5, 0.5)$. (d) $(\lambda, \theta_0, \theta_1, \theta_2) = (1, 0.2, 0.8, 0.6)$.

Theorem 1. *For* $l = 1, \cdots, n,$ $\delta_{0l} = I(T_{1l} = T_{2l}),$ $\delta_{1l} = I(T_{1l} < T_{2l}),$ *and* $\delta_{2l} = I(T_{1l} > T_{2l}),$ *We have*

$$(\delta_{0l}, \delta_{1l}, \delta_{2l}) \sim \text{Multinomial}\left(1; \frac{\theta_0}{\theta_0 + \theta_1 + \theta_2}, \frac{\theta_1}{\theta_0 + \theta_1 + \theta_2}, \frac{\theta_2}{\theta_0 + \theta_1 + \theta_2}\right), l = 1, \cdots, n. \quad (3)$$

Proof. For $l = 1, \cdots, n$, we have $\delta_{0l} + \delta_{1l} + \delta_{2l} = 1$,

$$P(T_1 < T_2) = \int_0^\infty \int_0^{t_2} f(t_1; \lambda, \theta_1) f(t_2; \lambda, \theta_0 + \theta_2) dt_1 dt_2 = \frac{\theta_1}{\theta_0 + \theta_1 + \theta_2},$$

$$P(T_1 > T_2) = \int_0^\infty \int_0^{t_1} f(t_1; \lambda, \theta_0 + \theta_1) f(t_2; \lambda, \theta_2) dt_2 dt_1 = \frac{\theta_2}{\theta_0 + \theta_1 + \theta_2}, \quad (4)$$

$$P(T_1 = T_2) = 1 - \frac{\theta_1}{\theta_0 + \theta_1 + \theta_2} - \frac{\theta_2}{\theta_0 + \theta_1 + \theta_2} = \frac{\theta_0}{\theta_0 + \theta_1 + \theta_2}.$$

Therefore, $(\delta_{0l}, \delta_{1l}, \delta_{2l}) \sim \text{Multinomial}(1; \theta_0/(\theta_0 + \theta_1 + \theta_2), \theta_1/(\theta_0 + \theta_1 + \theta_2), \theta_2/(\theta_0 + \theta_1 + \theta_2))$.

The likelihood function is

$$L(\lambda, \theta_0, \theta_1, \theta_2) = \prod_l \left[f_{T_1, T_2}(x_l, x_l; \lambda, \theta_0, \theta_1, \theta_2)\right]^{\delta_{0l}} \left[-\frac{\partial S_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)}{\partial t_1}\Big|_{(x_l, x_l)}\right]^{\delta_{1l}},$$

$$\times \left[-\frac{\partial S_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)}{\partial t_2}\Big|_{(x_l, x_l)}\right]^{\delta_{2l}}, \quad (5)$$

where

$$f_{T_1, T_2}(x_l, x_l; \lambda, \theta_0, \theta_1, \theta_2) = \frac{\theta_0}{\theta_0 + \theta_1 + \theta_2} f(t; \lambda, \theta_0 + \theta_1 + \theta_2),$$

$$= \theta_0 \exp\left\{\lambda x_l - \frac{\theta_0 + \theta_1 + \theta_2}{\lambda}\left(e^{\lambda x_l} - 1\right)\right\},$$

$$-\frac{\partial S_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)}{\partial t_1}\Big|_{(x_l, x_l)},$$

$$= \theta_1 \exp\left\{\lambda x_l - \frac{\theta_0 + \theta_1 + \theta_2}{\lambda}\left(e^{\lambda x_l} - 1\right)\right\},\tag{6}$$

$$-\frac{\partial S_{T_1, T_2}(t_1, t_2; \lambda, \theta_0, \theta_1, \theta_2)}{\partial t_2}\Big|_{(x_l, x_l)},$$

$$= \theta_2 \exp\left\{\lambda x_l - \frac{\theta_0 + \theta_1 + \theta_2}{\lambda}\left(e^{\lambda x_l} - 1\right)\right\}.$$

Then, we obtain

$$L(x; \lambda, \theta_0, \theta_1, \theta_2) = \theta_0^{n_0}\theta_1^{n_1}\theta_2^{n_2} \exp\left\{\lambda \sum_l x_l - \frac{\theta_0 + \theta_1 + \theta_2}{\lambda}\left[\sum_l\left(e^{\lambda x_l} - 1\right)\right]\right\}.\tag{7}$$

□

3. Classical Inference

3.1. Maximum Likelihood Estimates (MLEs). The MLEs of θ_0, θ_1, θ_2, and λ can be obtained by maximizing the logarithm of $L(x; \lambda, \theta_0, \theta_1, \theta_2)$. Set the first partial derivation of $\log L(x; \lambda, \theta_0, \theta_1, \theta_2)$ about θ_0, θ_1, θ_2, λ to 0, i.e.,

$$\frac{\partial \log L(x; \lambda, \theta_0, \theta_1, \theta_2)}{\partial \lambda} = \sum_l x_l + \frac{\theta_0 + \theta_1 + \theta_2}{\lambda^2}\left[\sum_l\left(e^{\lambda x_l} - 1\right)\right] - \frac{\theta_0 + \theta_1 + \theta_2}{\lambda}\left(\sum_l x_l e^{\lambda x_l}\right) = 0,$$

$$\frac{\partial \log L(x; \lambda, \theta_0, \theta_1, \theta_2)}{\partial \theta_j} = \frac{n_j}{\theta_j} - \frac{1}{\lambda}\left[\sum_l\left(e^{\lambda x_l} - 1\right)\right] = 0, \quad j = 0, 1, 2.\tag{8}$$

From (8), we get the MLEs of θ_0, θ_1, and θ_2 as

$$\widehat{\theta}_j(\lambda) = \frac{n_j \lambda}{\left[\sum_l\left(e^{\lambda x_l} - 1\right)\right]}, \quad j = 0, 1, 2.\tag{9}$$

Substituting $\widehat{\theta}_j(\lambda)$ into $\log L(x; \lambda, \theta_0, \theta_1, \theta_2)$, we obtain

$$h(\lambda) = \lambda \sum_l x_l + \sum_{j=0}^{2} \ln\left(\frac{n_j \lambda}{\sum_l\left(e^{\lambda x_l} - 1\right)}\right)^{n_j},\tag{10}$$

which is the profile logarithm likelihood function of λ.

We can show that $\partial^2 h(\lambda)/\partial \lambda^2 < 0$, which implies that $h(\lambda)$ is concave. Some iterative schemes can be used to find the MLE for λ, such as Newton–Raphson algorithm.

3.2. Bootstrap Confidence Intervals (CIs). Since it is hard to construct the exact CIs for the unknown parameters, we consider the Bootstrap method to construct CIs for parameters θ_0, θ_1, θ_2, and λ. The Bootstrap method is a resampling method to estimate some statistical characteristics for the unknown parameters by taking samples from

the original samples repeatedly; the obtained samples are called Bootstrap samples. This method has a great practical value since it does not need to assume the overall distribution or construct the pivot quantity. We generate the Bootstrap sample by the following three steps:

Step 1: for the fixed value of n and observed data $(x_1, x_2, \cdots, x_n)$, we get the estimates $\widehat{\lambda}$, $\widehat{\theta}_0$, $\widehat{\theta}_1$, and $\widehat{\theta}_2$ based on the maximum likelihood method.

Step 2: for the values of n, $\widehat{\lambda}$, $\widehat{\theta}_0$, $\widehat{\theta}_1$, and $\widehat{\theta}_2$, we generate the sample $(x_1^*, x_2^*, \cdots, x_n^*)$. Then, get the MLEs $\widehat{\lambda}'$, $\widehat{\theta}_0'$, $\widehat{\theta}_1'$, and $\widehat{\theta}_2'$.

Step 3: repeat Step 2 M times to obtain M sets of the values $\widehat{\lambda}'$, $\widehat{\theta}_0'$, $\widehat{\theta}_1'$, and $\widehat{\theta}_2'$. Arrange them as follows to get the Bootstrap sample:

$$\left\{ \widehat{\theta}_{0[1]}' < \cdots < \widehat{\theta}_{0[M]}', \widehat{\theta}_{1[1]}' < \cdots < \widehat{\theta}_{1[M]}', \widehat{\theta}_{2[1]}' < \cdots < \widehat{\theta}_{2[M]}', \widehat{\lambda}_{[1]}' < \cdots < \widehat{\lambda}_{[M]}' \right\}. \tag{11}$$

Based on the Bootstrap sample and by percentile Bootstrap (Boot-P) method, we construct the Boot-P CIs for θ_0, θ_1, θ_2, λ at $1 - \gamma$ confidence level as

$$\left(\widehat{\theta}_{0[M\gamma/2]}'', \widehat{\theta}_{0[M(1-\gamma/2)]}'' \right), \left(\widehat{\theta}_{1[M\gamma/2]}'', \widehat{\theta}_{1[M(1-\gamma/2)]}'' \right), \left(\widehat{\theta}_{2[M\gamma/2]}'', \widehat{\theta}_{2[M(1-\gamma/2)]}'' \right), \left(\lambda_{[M\gamma/2]}'', \lambda_{[M(1-\gamma/2)]}'' \right). \tag{12}$$

4. Bayesian Inference and HPD CIs

4.1. Conjugate Prior.
In this section, we suppose the shape parameter λ is known. Denote $\theta = \theta_0 + \theta_1 + \theta_2$, which has a Gamma prior with hyperparameters a and b as

$$\pi(\theta) = \left(\frac{b^a}{\Gamma(a)} \right) \theta^{a-1} e^{-b\theta}, \tag{13}$$

$$a > 0, b > 0, \theta > 0.$$

Due to $\theta_0/\theta + \theta_1/\theta + \theta_2/\theta = 1$, so given θ, $(\theta_1/\theta, \theta_2/\theta)$ follows a Dirichlet prior with hyper parameters c_0, c_1, and c_2, that is,

$$\pi_D\left(\frac{\theta_1}{\theta}, \frac{\theta_2}{\theta} \Big| \theta \right) = \frac{\Gamma\left(\sum_{i=0}^2 c_i \right)}{\prod_{i=0}^2 \Gamma(c_i)} \prod_{i=0}^2 \left(\frac{\theta_i}{\theta} \right)^{c_i - 1}, \quad \theta_i > 0, c_i > 0, i = 0, 1, 2. \tag{14}$$

Therefore, the joint prior of θ_0, θ_1, and θ_2 becomes

$$\pi_1(\theta_0, \theta_1, \theta_2; a, b, c_0, c_1, c_2) = \frac{\Gamma(c)}{\Gamma(a)} (b\theta)^{a-c} \prod_{i=0}^2 \frac{b^{c_i}}{\Gamma(c_i)} \theta_i^{c_i - 1} \exp(-b\theta_i), \tag{15}$$

where $c = c_0 + c_1 + c_2$.

4.2. Jeffreys Prior.
According to Jeffreys [20], Jeffreys prior is proportional to the square root of the determinant of the Fisher information matrix. From (7), we obtain the Fisher information matrix of $(\theta_0, \theta_1, \theta_2)$ as

$$I = I(\theta_0, \theta_1, \theta_2) = \begin{pmatrix} \dfrac{n_0}{\theta_0^2} & 0 & 0 \\ 0 & \dfrac{n_1}{\theta_1^2} & 0 \\ 0 & 0 & \dfrac{n_2}{\theta_2^2} \end{pmatrix}. \tag{16}$$

From Theorem 1, we have $n_i = n \cdot \theta_i/(\theta_0 + \theta_1 + \theta_2)$, $i = 0, 1, 2$, so $I(\theta_0, \theta_1, \theta_2)$ can be written as

$$I = n \begin{pmatrix} \dfrac{1}{(\theta_0 \theta)} & 0 & 0 \\[2ex] 0 & \dfrac{1}{(\theta_1 \theta)} & 0 \\[3ex] 0 & 0 & \dfrac{1}{(\theta_2 \theta)} \end{pmatrix}. \tag{17}$$

Thus, the Jeffreys prior is given by

$$\pi_2(\theta_0, \theta_1, \theta_2) \propto \sqrt{\frac{1}{\theta_0 \theta_1 \theta_2 \theta^3}}. \tag{18}$$

Theorem 2. *Based on the Jeffreys prior $\pi_2(\theta_0, \theta_1, \theta_2)$, the joint posterior distribution of $(\theta_0, \theta_1, \theta_2)$ is proper.*

Proof. From (6) and (7), we obtain the joint posterior distribution of $(\theta_0, \theta_1, \theta_2)$ based on $\pi_2(\theta_0, \theta_1, \theta_2)$ as

$$\pi_2(\theta_0, \theta_1, \theta_2 | x) = \frac{L(x; \lambda, \theta_0, \theta_1, \theta_2) \pi_2(\theta_0, \theta_1, \theta_2)}{\int_0^\infty \int_0^\infty \int_0^\infty L(x; \lambda, \theta_0, \theta_1, \theta_2) \pi_2(\theta_0, \theta_1, \theta_2) d\theta_0 d\theta_1 d\theta_2},$$

$$\propto \theta^{(-3/2)} \theta_0^{(n_0 - 1/2)} \theta_1^{(n_1 - 1/2)} \theta_2^{(n_2 - 1/2)} e^{(-A\theta/\lambda)}. \tag{19}$$

Integrating $\pi_2(\theta_0, \theta_1, \theta_2 | x)$ with respect to θ_0, θ_1, and θ_2, we obtain

$$\int_0^\infty \int_0^\infty \int_0^\infty \theta^{-3/2} \theta_0^{n_0 - 1/2} \theta_1^{n_1 - 1/2} \theta_2^{n_2 - 1/2} e^{(-A\theta/\lambda)} d\theta_0 d\theta_1 d\theta_2,$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \left(\frac{\theta_0}{\theta}\right)^{n_0 - 1/2} \left(\frac{\theta_1}{\theta}\right)^{n_1 - 1/2} \left(\frac{\theta_2}{\theta}\right)^{n_2 - 1/2} \theta^{n-3} \exp\left\{-\left(\frac{A}{\lambda}\right)\theta\right\} d\theta_0 d\theta_1 d\theta_2,$$

$$= \int_{0 < \theta_0/\theta + \theta_1/\theta < 1} \prod_{i=0}^1 \left(\frac{\theta_i}{\theta}\right)^{n_i - 1/2} \left(1 - \sum_{i=0}^1 \frac{\theta_i}{\theta}\right)^{n_2 - 1/2} d\frac{\theta_0}{\theta} d\frac{\theta_1}{\theta} \cdot \int_0^\infty \theta^{n-1} \exp\left\{-\frac{A}{\lambda}\theta\right\} d\theta, \tag{20}$$

$$= B\left(n_0 + \frac{1}{2}, n_1 + n_2 + 1\right) B\left(n_1 + \frac{1}{2}, n_2 + \frac{1}{2}\right) \frac{\Gamma(n)}{(A/\lambda)^n} < \infty,$$

where $A = \sum_{i=1}^n (e^{\lambda x_i} - 1)$ and $B(\cdot, \cdot)$ is a beta function.

Thus, the joint posterior distribution of $(\theta_0, \theta_1, \theta_2)$ based on $\pi_2(\theta_0, \theta_1, \theta_2)$ is proper. □

4.3. Reference Priors. Bernardo [18] and Berger and Bernardo [19] proposed the reference prior which plays a vital role in the objective Bayesian inference. We set $\mu_0 \equiv \theta = \theta_0 + \theta_1 + \theta_2$, $\mu_1 = \theta_0/\theta$, and $\mu_2 = \theta_1/\theta$; the transformation from $(\theta_0, \theta_1, \theta_2)$ to (μ_0, μ_1, μ_2) is one-to-one with the inverse transformation $\theta_0 = \mu_0 \mu_1$, $\theta_1 = \mu_0 \mu_2$, and

$\theta_2 = \mu_0(1 - \mu_1 - \mu_2)$. The Jacobian matrix of the transformation has the form

$$J = \begin{pmatrix} \mu_1 & \mu_0 & 0 \\ \mu_2 & 0 & \mu_0 \\ 1 - \mu_1 - \mu_2 & -\mu_0 & -\mu_0 \end{pmatrix}, 0 < \mu_0 < \infty, 0 < \mu_1 + \mu_2 < 1. \tag{21}$$

The likelihood function (3) becomes

$$L(x; \lambda, \mu_0, \mu_1, \mu_2) = \mu_1^{n_0} \mu_2^{n_1} (1 - \mu_1 - \mu_2)^{n_2} \mu_0^n \exp\left\{\lambda \sum_l x_l - \left(\frac{\mu_0}{\lambda}\right)\left[\sum_l (e^{\lambda x_l} - 1)\right]\right\}. \tag{22}$$

The Fisher information matrix of (μ_0, μ_1, μ_2) can be written as

$$I_1 = J'IJ = n \begin{pmatrix} 1/\mu_0^2 & 0 & 0 \\ & & \\ 0 & \dfrac{1}{\mu_1} + \dfrac{1}{(1-\mu_1-\mu_2)} & \dfrac{1}{(1-\mu_1-\mu_2)} \\ & & \\ 0 & \dfrac{1}{(1-\mu_1-\mu_2)} & \dfrac{1}{\mu_2} + \dfrac{1}{(1-\mu_1-\mu_2)} \end{pmatrix}. \tag{23}$$

Theorem 3.

(i) *Under the ordering groups* $\{\mu_0, (\mu_1, \mu_2)\}$ *and* $\{(\mu_1, \mu_2), \mu_0\}$, *the reference priors are the same, which is given by* $\omega_{R_1}(\mu_0, \mu_1, \mu_2) = \sqrt{1/\mu_0^2\mu_1\mu_2(1-\mu_1-\mu_2)}$; *the corresponding reference prior for* $(\theta_0, \theta_1, \theta_2)$ *is* $\pi_2(\theta_0, \theta_1, \theta_2) = \sqrt{1/\theta_0\theta_1\theta_2\theta^3}$

(ii) *Under the ordering groups* $\{\mu_0, \mu_1, \mu_2\}$, $\{\mu_0, \mu_2, \mu_1\}$, $\{\mu_1, \mu_0, \mu_2\}$, *and* $\{\mu_1, \mu_2, \mu_0\}$, *the reference priors are the same, which is given by* $\omega_{R_2}(\mu_0, \mu_1, \mu_2) = \sqrt{1/\mu_0^2\mu_1\mu_2(1-\mu_1-\mu_2)(1-\mu_1)}$; *the corresponding reference prior for* $(\theta_0, \theta_1, \theta_2)$ *is* $\pi_3(\theta_0, \theta_1, \theta_2) = \sqrt{1/\theta^2\theta_0\theta_1\theta_2(\theta_1+\theta_2)}$

(iii) *Under the ordering groups* $\{\mu_2, \mu_0, \mu_1\}$ *and* $\{\mu_2, \mu_1, \mu_0\}$, *the reference priors are the same, which is given by* $\omega_{R_3}(\mu_0, \mu_1, \mu_2) = \sqrt{1/\mu_0^2\mu_1\mu_2(1-\mu_1-\mu_2)(1-\mu_2)}$; *the corresponding reference prior for* $(\theta_0, \theta_1, \theta_2)$ *is* $\pi_4(\theta_0, \theta_1, \theta_2)) = \sqrt{1/\theta^2\theta_0\theta_1\theta_2(\theta_0+\theta_2)}$

Proof.

(i) The Fisher information matrix of (μ_0, μ_1, μ_2) is

$$I_1 = \begin{pmatrix} \sum_{11} & 0 \\ 0 & \sum_{22} \end{pmatrix}, \tag{24}$$

where $\sum_{11} = n/\mu_0^2$ and $\sum_{22} = n$ $\begin{pmatrix} 1/\mu_1 + 1/(1-\mu_1-\mu_2) & 1/(1-\mu_1-\mu_2) \\ 1/(1-\mu_1-\mu_2) & 1/\mu_2 + 1/(1-\mu_1-\mu_2) \end{pmatrix}$.
The reference prior for the ordering groups $\{\mu_0, (\mu_1, \mu_2)\}$ and $\{(\mu_1, \mu_2), \mu_0\}$ is the same as in [21], which is given by

$$\omega_{R_1}(\mu_0, \mu_1, \mu_2) \propto \left|\sum_{11}\right|^{1/2} \left|\sum_{22}\right|^{1/2} \propto \sqrt{\dfrac{1}{\mu_0^2\mu_1\mu_2(1-\mu_1-\mu_2)}}. \tag{25}$$

(ii) The inverse of I_1 is

$$H = \dfrac{1}{n} \begin{pmatrix} \mu_0^2 & 0 & 0 \\ 0 & \mu_1(1-\mu_1) & -\mu_1\mu_2 \\ 0 & -\mu_1\mu_2 & \mu_2(1-\mu_2) \end{pmatrix}. \tag{26}$$

(iii) According the notations in [18], we obtain $h_1 = 1/\mu_0^2$, $h_2 = 1/\mu_1(1-\mu_1)$, and $h_3 = (1-\mu_1)/(\mu_2(1-\mu_1-\mu_2))$.

Choose the compact sets $\Omega_k = \{(\mu_0, \mu_1, \mu_2)| a_{0k} < \mu_0 < b_{0k}, a_{1k} < \mu_1, a_{2k} < \mu_2, \mu_1 + \mu_2 < d_k\}$, such that $a_{0k}, a_{1k}, a_{2k} \longrightarrow 0$, $b_{0k} \longrightarrow \infty$, and $d_k \longrightarrow 1$, as $k \longrightarrow \infty$. Then, we have

$$\pi^k(\mu_0, \mu_1, \mu_2) = \dfrac{\sqrt{|h_1|}\sqrt{|h_2|}\sqrt{|h_3|}}{\int_{a_{0k}}^{b_{0k}} \sqrt{|h_1|}d\mu_0 \cdot \int_{a_{1k}}^{d_k-\mu_2^0} \sqrt{|h_2|}d\mu_1 \cdot \int_{a_{2k}}^{d_k-\mu_1} \sqrt{|h_3|}d\mu_2} I_{\Omega_k}(\mu_0, \mu_1, \mu_2), \tag{27}$$

where $\int_{a_{0k}}^{b_{0k}} \sqrt{|h_1|}d\mu_0 = \int_{a_{0k}}^{b_{0k}} 1/\mu_0 d\mu_0 = \log b_{0k} - \log a_{0k}$:

$$\int_{a_{1k}}^{d_k-\mu_2^0} \sqrt{|h_2|}d\mu_1 = \int_{a_{1k}}^{d_k-\mu_2^0} \sqrt{\dfrac{1}{\mu_1(1-\mu_1)}}d\mu_1 = -\arcsin\left(1 - 2(d_k - \mu_2^0)\right) + \arcsin(1 - 2a_{1k}),$$

$$\int_{a_{2k}}^{d_k-\mu_1} \sqrt{|h_3|}d\mu_2 = \int_{a_{2k}}^{d_k-\mu_1} \sqrt{\dfrac{1-\mu_1}{\mu_2(1-\mu_1-\mu_2)}}d\mu_2, \tag{28}$$

$$= (1-\mu_1)^{1/2}\left(-\arcsin\left(\dfrac{1-\mu_1-2(d_k-\mu_1)}{1-\mu_1}\right) + \arcsin\left(\dfrac{1-\mu_1-2a_{2k}}{1-\mu_1}\right)\right).$$

Then, we get the reference prior as

$$\omega_{R_2}(\mu_0, \mu_1, \mu_2) = \lim_{k \to \infty} \frac{\pi^k(\mu_0, \mu_1, \mu_2)}{\pi^k(\mu_0^*, \mu_1^*, \mu_2^*)} \propto \sqrt{\frac{1}{\mu_0^2 \mu_1 \mu_2 (1 - \mu_1 - \mu_2)(1 - \mu_1)}}, \tag{29}$$

where $(\mu_0^*, \mu_1^*, \mu_2^*)$ is an inner point of Ω_k.

Similarly, under the ordering group $\{\mu_0, \mu_2, \mu_1\}$, the reference prior is $\omega_{R_2}(\mu_0, \mu_1, \mu_2)$.

The Fisher information matrix of $\{\mu_1, \mu_0, \mu_2\}$ is

$$I_2 = n \begin{pmatrix} \dfrac{1}{\mu_1} + \dfrac{1}{(1 - \mu_1 - \mu_2)} & 0 & \dfrac{1}{(1 - \mu_1 - \mu_2)} \\ 0 & \dfrac{1}{\mu_0^2} & 0 \\ \dfrac{1}{(1 - \mu_1 - \mu_2)} & 0 & \dfrac{1}{\mu_2} + \dfrac{1}{(1 - \mu_1 - \mu_2)} \end{pmatrix}. \tag{30}$$

The inverse of I_2 is

$$H_1 = \frac{1}{n} \begin{pmatrix} \mu_1 (1 - \mu_1) & 0 & -\mu_1 \mu_2 \\ 0 & \mu_0^2 & 0 \\ -\mu_1 \mu_2 & 0 & \mu_2 (1 - \mu_2) \end{pmatrix}. \tag{31}$$

Similarly, we obtain $h_1 = 1/\mu_1 (1 - \mu_1)$, $h_2 = 1/\mu_0^2$, and $h_3 = (1 - \mu_1)/(\mu_2 (1 - \mu_1 - \mu_2))$.

Choose the compact sets $\Omega_k = \{(\mu_1, \mu_0, \mu_2) \mid a_{0k} < \mu_1, \ a_{1k} < \mu_0 < b_{1k}, a_{2k} < \mu_2, \mu_1 + \mu_2 < d_k\}$, such that $a_{0k}, a_{1k}, a_{2k} \to 0$, $b_{1k} \to \infty$, and $d_k \to 1$, as $k \to \infty$. Then, we have

$$\pi^k(\mu_1, \mu_0, \mu_2) = \frac{\sqrt{|h_1|}\sqrt{|h_2|}\sqrt{|h_3|}}{\int_{a_{0k}}^{d_k - u_2^0} \sqrt{|h_1|} d\mu_1 \cdot \int_{a_{1k}}^{b_{1k}} \sqrt{|h_2|} d\mu_0 \cdot \int_{a_{2k}}^{d_k - \mu_1} \sqrt{|h_3|} d\mu_2} I_{\Omega_k}(\mu_1, \mu_0, \mu_2), \tag{32}$$

where $\int_{a_{0k}}^{d_k - u_2^0} \sqrt{|h_1|} d\mu_1 = \int_{a_{0k}}^{d_k - u_2^0} \sqrt{1/\mu_1 (1 - \mu_1)} d\mu_1 = -\arcsin(1 - 2(d_k - u_2^0)) + \arcsin(1 - 2a_{0k})$,

$$\int_{a_{1k}}^{b_{1k}} \sqrt{|h_2|} d\mu_0 = \int_{a_{1k}}^{b_{1k}} \frac{1}{\mu_0} d\mu_0 = \log b_{1k} - \log a_{1k},$$

$$\int_{a_{2k}}^{d_k - \mu_1} \sqrt{|h_3|} d\mu_2 = \int_{a_{2k}}^{d_k - \mu_1} \sqrt{\frac{1 - \mu_1}{\mu_2 (1 - \mu_1 - \mu_2)}} d\mu_2,$$

$$= (1 - \mu_1)^{1/2} \left(-\arcsin\left(\frac{1 - \mu_1 - 2(d_k - u_1)}{1 - \mu_1}\right) + \arcsin\left(\frac{1 - \mu_1 - 2a_{2k}}{1 - \mu_1}\right) \right). \tag{33}$$

Let $(\mu_1^*, \mu_0^*, \mu_2^*)$ be an inner point of Ω_k; we get the reference prior as

$$\omega_{R_2}(\mu_0, \mu_1, \mu_2) = \lim_{k \to \infty} \frac{\pi^k(\mu_1, \mu_0, \mu_2)}{\pi^k(\mu_1^*, \mu_0^*, \mu_2^*)} \propto \sqrt{\frac{1}{\mu_0^2 \mu_1 \mu_2 (1 - \mu_1 - \mu_2)(1 - \mu_1)}}. \tag{34}$$

Similarly, under the ordering group $\{\mu_1, \mu_2, \mu_0\}$, the reference prior is $\omega_{R_2}(\mu_0, \mu_1, \mu_2)$.

(v) The Fisher information matrix of $\{\mu_2, \mu_1, \mu_0\}$ is

$$
I_3 = n \begin{pmatrix} \dfrac{1}{\mu_2} + \dfrac{1}{(1 - \mu_1 - \mu_2)} & \dfrac{1}{(1 - \mu_1 - \mu_2)} & 0 \\[3mm] \dfrac{1}{(1 - \mu_1 - \mu_2)} & \dfrac{1}{\mu_1} + \dfrac{1}{(1 - \mu_1 - \mu_2)} & 0 \\[3mm] 0 & 0 & \dfrac{1}{\mu_0^2} \end{pmatrix}. \tag{35}
$$

The inverse of I_3 is

$$
H_2 = \frac{1}{n} \begin{pmatrix} \mu_2(1 - \mu_2) & -\mu_1\mu_2 & 0 \\[2mm] -\mu_1\mu_2 & \mu_1(1 - \mu_1) & 0 \\[2mm] 0 & 0 & \mu_0^2 \end{pmatrix}. \tag{36}
$$

Then, we obtain $h_1 = 1/\mu_2(1 - \mu_2)$, $h_2 = (1 - \mu_2)/(\mu_1(1 - \mu_1 - \mu_2))$, and $h_3 = 1/\mu_0^2$.

Choose the compact sets $\Omega_k = \{(\mu_2, \mu_1, \mu_0) | a_{0k} < \mu_2, a_{1k} < \mu_1, \mu_2 + \mu_1 < d_k, a_{2k} < \mu_0 < b_{2k}\}$, such that $a_{0k}, a_{1k}, a_{2k} \longrightarrow 0$, $b_{2k} \longrightarrow \infty$, and $d_k \longrightarrow 1$, as $k \longrightarrow \infty$. Then, we have

$$
\pi^k(\mu_2, \mu_1, \mu_0) = \frac{\sqrt{|h_1|}\sqrt{|h_2|}\sqrt{|h_3|}}{\int_{a_{0k}}^{d_k - u_1^0} \sqrt{|h_1|}d\mu_2 \cdot \int_{a_{1k}}^{d_k - \mu_2} \sqrt{|h_2|}d\mu_1 \cdot \int_{a_{2k}}^{b_{2k}} \sqrt{|h_3|}d\mu_0} I_{\Omega_k}(\mu_2, \mu_1, \mu_0), \tag{37}
$$

where $\int_{a_{0k}}^{d_k - u_1^0} \sqrt{|h_1|}d\mu_2 = \int_{a_{0k}}^{d_k - u_1^0} \sqrt{1/\mu_2(1 - \mu_2)}d\mu_2 = -\arcsin(1 - 2(d_k - u_1^0)) + \arcsin(1 - 2a_{0k})$,

$$
\int_{a_{1k}}^{d_k - \mu_2} \sqrt{|h_2|}d\mu_1 = \int_{a_{1k}}^{d_k - \mu_2} \sqrt{\frac{1 - \mu_2}{\mu_1(1 - \mu_1 - \mu_2)}}d\mu_1,
$$

$$
= (1 - \mu_2)^{1/2} \left(-\arcsin\left(\frac{1 - \mu_2 - 2(d_k - \mu_2)}{1 - \mu_2}\right) + \arcsin\left(\frac{1 - \mu_2 - 2a_{1k}}{1 - \mu_2}\right) \right), \tag{38}
$$

$$
\int_{a_{2k}}^{b_{2k}} \sqrt{|h_3|}d\mu_0 = \int_{a_{2k}}^{b_{2k}} \frac{1}{\mu_0}d\mu_0 = \log b_{2k} - \log a_{2k}.
$$

Let $(\mu_2^*, \mu_1^*, \mu_0^*)$ be an inner point of Ω_k, we obtain the reference prior as

$$
\omega_{R_3}(\mu_0, \mu_1, \mu_2) = \lim_{k \longrightarrow \infty} \frac{\pi^k(\mu_2, \mu_1, \mu_0)}{\pi^k(\mu_2^*, \mu_1^*, \mu_0^*)} \propto \sqrt{\frac{1}{\mu_0^2 \mu_1 \mu_2 (1 - \mu_1 - \mu_2)(1 - \mu_2)}}. \tag{39}
$$

Similarly, under the ordering group $\{\mu_2, \mu_0, \mu_1\}$, the reference prior is $\omega_{R_3}(\mu_0, \mu_1, \mu_2)$. According to the one-to-one transformation from (μ_0, μ_1, μ_2) to $(\theta_0, \theta_1, \theta_2)$, we can obtain the reference priors $\pi_2(\mu_0, \mu_1, \mu_2)$, $\pi_3(\mu_0, \mu_1, \mu_2)$, $\pi_4(\mu_0, \mu_1, \mu_2)$ from ω_{R_1}, ω_{R_2}, and ω_{R_3}, respectively. □

Theorem 4. *Based on the reference priors $\pi_3(\theta_0, \theta_1, \theta_2)$ and $\pi_4(\theta_0, \theta_1, \theta_2)$, the posterior distributions of $(\theta_0, \theta_1, \theta_2)$ are proper.*

Proof. The joint posterior distributions of $(\theta_0, \theta_1, \theta_2)$ based on reference prior $\pi_3(\theta_0, \theta_1, \theta_2)$ and $\pi_4(\theta_0, \theta_1, \theta_2)$ are, respectively, as

$$\pi_3\left(\theta_0,\theta_1,\theta_2|x\right) = \frac{L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_3\left(\theta_0,\theta_1,\theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_3\left(\theta_0,\theta_1,\theta_2\right)d\theta_0 d\theta_1 d\theta_2},$$

$$\propto \theta^{-1}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2}\left(\theta_1+\theta_2\right)^{-1/2}\exp\left\{-\frac{A\theta}{\lambda}\right\},$$

$$\pi_4\left(\theta_0,\theta_1,\theta_2|x\right) = \frac{L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_4\left(\theta_0,\theta_1,\theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_4\left(\theta_0,\theta_1,\theta_2\right)d\theta_0 d\theta_1 d\theta_2},$$

$$\propto \theta^{-1}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2}\left(\theta_0+\theta_2\right)^{-1/2}\exp\left\{-\frac{A\theta}{\lambda}\right\}.$$

$$(40)$$

Integrating $\pi_3\left(\theta_0,\theta_1,\theta_2|x\right)$ and $\pi_4\left(\theta_0,\theta_1,\theta_2|x\right)$ with respect to $\theta_0,\theta_1,$ and θ_2, respectively, we obtain

$$\int_0^\infty \int_0^\infty \int_0^\infty \theta^{-1}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2}\left(\theta_1+\theta_2\right)^{-1/2}\exp\left\{-\frac{A\theta}{\lambda}\right\}d\theta_0 d\theta_1 d\theta_2,$$

$$= B\left(n_0+\frac{1}{2},n_1+n_2+\frac{1}{2}\right)B\left(n_1+\frac{1}{2},n_2+\frac{1}{2}\right)\frac{\Gamma(n)}{(A/\lambda)^n}<\infty,$$

$$\int_0^\infty \int_0^\infty \int_0^\infty \theta^{-1}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2}\left(\theta_0+\theta_2\right)^{-1/2}\exp\left\{-\frac{A\theta}{\lambda}\right\}d\theta_0 d\theta_1 d\theta_2,$$

$$= B\left(n_1+\frac{1}{2},n_0+n_2+\frac{1}{2}\right)B\left(n_0+\frac{1}{2},n_2+\frac{1}{2}\right)\frac{\Gamma(n)}{(A/\lambda)^n}<\infty.$$

$$(41)$$

Thus, the posterior distributions of $(\theta_0,\theta_1,\theta_2)$ based on $\pi_3\left(\theta_0,\theta_1,\theta_2\right)$ and $\pi_4\left(\theta_0,\theta_1,\theta_2\right)$ are proper.   □

4.4. Bayesian Estimates. The joint posterior distributions of $(\theta_0,\theta_1,\theta_2)$ based on $\pi_1,\pi_2,\pi_3,$ and π_4 are, respectively, as

$$\pi_1\left(\theta_0,\theta_1,\theta_2|x\right) = \frac{L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_1\left(\theta_0,\theta_1,\theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_1\left(\theta_0,\theta_1,\theta_2\right)d\theta_0 d\theta_1 d\theta_2},$$

$$(42)$$

where

$$\int_0^\infty \int_0^\infty \int_0^\infty L\left(x;\lambda,\theta_0,\theta_1,\theta_2\right)\pi_1\left(\theta_0,\theta_1,\theta_2\right)d\theta_0 d\theta_1 d\theta_2$$

$$= w_1 w_2 \exp\left\{\lambda\sum_l x_l\right\}\int_{0<\theta_0/\theta+\theta_1/\theta<1}\prod_{i=0}^1\left(\frac{\theta_i}{\theta}\right)^{n_i+c_i-1}\left(1-\sum_{i=0}^1\frac{\theta_i}{\theta}\right)^{n_2+c_2-1}d\frac{\theta_0}{\theta}d\frac{\theta_1}{\theta},\int_0^\infty \theta^{n+a+1}e^{-(A/\lambda+b)\theta}d\theta$$

$$= w_1 w_2 \exp\left\{\lambda\sum_l x_l\right\}B\left(n_0+c_0,n_1+c_1+n_2+c_2\right)B\left(n_1+c_1,n_2+c_2\right)\frac{\Gamma(n+a+2)}{(A/\lambda+b)^{n+a+2}},$$

$$(43)$$

where $w_1 = \Gamma\left(\sum_{i=0}^{2} c_i\right) b^{a-c_0-c_1-c_2}/\Gamma(a)$ and $w_2 = \prod_{i=0}^{2} b^{c_i}/\Gamma(c_i)$.   Thus, we obtain

$$\pi_1\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{\theta^{a-c_0-c_1-c_2}\theta_0^{n_0+c_0-1}\theta_1^{n_1+c_1-1}\theta_2^{n_2+c_2-1} \exp\{-(A/\lambda + b)\theta\}}{B\left(n_0 + c_0, n_1 + c_1 + n_2 + c_2\right)B\left(n_1 + c_1, n_2 + c_2\right)\Gamma(n+a)/\left(A/\lambda + b\right)^{n+a}}. \tag{44}$$

Similarly,

$$\pi_2\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_2\left(\theta_0, \theta_1, \theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_2\left(\theta_0, \theta_1, \theta_2\right)d\theta_0 d\theta_1 d\theta_2}, \tag{45}$$

where

$$\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_2\left(\theta_0, \theta_1, \theta_2\right)d\theta_0 d\theta_1 d\theta_2,$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \theta^{-3/2}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2} \exp\left\{\lambda \sum_l x_l - \frac{A}{\lambda}\theta\right\}d\theta_0 d\theta_1 d\theta_2, \tag{46}$$

$$= \exp\left\{\lambda \sum_l x_l\right\}B\left(n_0 + \frac{1}{2}, n_1 + n_2 + 1\right)B\left(n_1 + \frac{1}{2}, n_2 + \frac{1}{2}\right)\frac{\Gamma(n)}{(A/\lambda)^n}.$$

We obtain

$$\pi_2\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{\theta^{-3/2}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2} \exp\{-A\theta/\lambda\}}{B\left(n_0 + 1/2, n_1 + n_2 + 1\right)B\left(n_1 + 1/2, n_2 + 1/2\right)\Gamma(n)/\left(A/\lambda\right)^n},$$

$$\pi_3\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_3\left(\theta_0, \theta_1, \theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_3\left(\theta_0, \theta_1, \theta_2\right)d\theta_0 d\theta_1 d\theta_2}, \tag{47}$$

where

$$\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right)\pi_3\left(\theta_0, \theta_1, \theta_2\right)d\theta_0 d\theta_1 d\theta_2,$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \theta^{-1}\theta_0^{n_0-1/2}\theta_1^{n_1-1/2}\theta_2^{n_2-1/2}\left(\theta_1 + \theta_2\right)^{-1/2} \exp\left\{\lambda \sum_l x_l - \frac{A\theta}{\lambda}\right\}d\theta_0 d\theta_1 d\theta_2, \tag{48}$$

$$= \exp\left\{\lambda \sum_l x_l\right\}B\left(n_0 + \frac{1}{2}, n_1 + n_2 + \frac{1}{2}\right)B\left(n_1 + \frac{1}{2}, n_2 + \frac{1}{2}\right)\frac{\Gamma(n)}{(A/\lambda)^n}.$$

We obtain

$$\pi_3\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{\theta^{-1} \theta_0^{n_0 - 1/2} \theta_1^{n_1 - 1/2} \theta_2^{n_2 - 1/2} \left(\theta_1 + \theta_2\right)^{-1/2} \exp\{-A\theta/\lambda\}}{B\left(n_0 + 1/2, n_1 + n_2 + 1/2\right) B\left(n_1 + 1/2, n_2 + 1/2\right) \Gamma(n) / (A/\lambda)^n},$$

$$\pi_4\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right) \pi_4\left(\theta_0, \theta_1, \theta_2\right)}{\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right) \pi_4\left(\theta_0, \theta_1, \theta_2\right) d\theta_0 d\theta_1 d\theta_2},$$

(49)

where

$$\int_0^\infty \int_0^\infty \int_0^\infty L\left(x; \lambda, \theta_0, \theta_1, \theta_2\right) \pi_4\left(\theta_0, \theta_1, \theta_2\right) d\theta_0 d\theta_1 d\theta_2,$$

$$= \int_0^\infty \int_0^\infty \int_0^\infty \theta^{-1} \theta_0^{n_0 - 1/2} \theta_1^{n_1 - 1/2} \theta_2^{n_2 - 1/2} \left(\theta_0 + \theta_2\right)^{-1/2} \exp\left\{\lambda \sum_l x_l - \frac{A\theta}{\lambda}\right\} d\theta_0 d\theta_1 d\theta_2,$$

$$= \exp\left\{\lambda \sum_l x_l\right\} B\left(n_1 + \frac{1}{2}, n_0 + n_2 + \frac{1}{2}\right) B\left(n_0 + \frac{1}{2}, n_2 + \frac{1}{2}\right) \frac{\Gamma(n)}{(A/\lambda)^n}.$$

(50)

Then, we have

$$\pi_4\left(\theta_0, \theta_1, \theta_2 | x\right) = \frac{\theta^{-1} \theta_0^{n_0 - 1/2} \theta_1^{n_1 - 1/2} \theta_2^{n_2 - 1/2} \left(\theta_0 + \theta_2\right)^{-1/2} \exp\{-A\theta/\lambda\}}{B\left(n_1 + 1/2, n_0 + n_2 + 1/2\right) B\left(n_0 + 1/2, n_2 + 1/2\right) \Gamma(n) / (A/\lambda)^n}.$$

(51)

From (9)–(12), we get the Bayesian estimates of parameters θ_0, θ_1, θ_2, and θ against squared error loss function based on π_1, π_2, π_3, and π_4, respectively, which are listed in Table 1.

4.5. HPD Credible Intervals. The HPD credible intervals of parameters θ_0, θ_1, θ_2, and θ can be constructed by the Monte Carlo method studied by Chen and Shao [22].

Step 1: given the value of n and the observed data $(x_1, x_2, \cdots, x_n)$, compute the Bayesian estimates of $\widehat{\theta}_0$, $\widehat{\theta}_1$, $\widehat{\theta}_2$, and $\widehat{\theta}$ based on π_1, π_2, π_3, and π_4, respectively.

Step 2: repeat Step 1 M times; we obtain M sets of the values $\widehat{\theta}_0$, $\widehat{\theta}_1$, $\widehat{\theta}_2$, and $\widehat{\theta}$ based on π_1, π_2, π_3, and π_4, respectively. Arrange them in the ascending order, we obtain

$$\widehat{\theta}_{0\pi_k[1]} < \cdots < \widehat{\theta}_{0\pi_k[M]}, \widehat{\theta}_{1\pi_k[1]} < \cdots < \widehat{\theta}_{1\pi_k[M]}, \widehat{\theta}_{2\pi_k[1]} < \cdots < \widehat{\theta}_{2\pi_k[M]}, \widehat{\theta}_{\pi_k[1]} < \cdots < \widehat{\theta}_{\pi_k[M]}, k = 1, 2, 3, 4.$$

(52)

Step 3: compute the CIs at $1 - \gamma$ confidence level as

$$\left(\widehat{\theta}_{\nu\pi_k[w]}, \widehat{\theta}_{\nu\pi_k[w+(1-\gamma)M]}\right), \left(\widehat{\theta}_{\pi_k[w]}, \widehat{\theta}_{\pi_k[w+(1-\gamma)M]}\right), \nu = 0, 1, 2; w = 1, 2, \cdots, M - (1-\gamma)M; k = 1, 2, 3, 4.$$

(53)

Step 4: the HPD CIs for θ_ν, $\nu = 0, 1, 2$, and θ are the shortest intervals among $\left(\widehat{\theta}_{\nu\pi_k[w]}, \widehat{\theta}_{\nu\pi_k[w+(1-\gamma)M]}\right)$, $\left(\widehat{\theta}_{\pi_k[w]}, \widehat{\theta}_{\pi_k[w+(1-\gamma)M]}\right)$, and $w = 1, 2, \cdots, M - (1-\gamma)M$, respectively.

Table 1: Bayesian estimates of parameters based on different priors.

Prior	θ_0	θ_1	θ_2	θ
π_1	$\lambda(n_0 + c_0)(n + a)/(n + c_0 + c_1 + c_2)(A + b\lambda)$	$\lambda(n_1 + c_1)(n + a)/(n + c_0 + c_1 + c_2)(A + b\lambda)$	$\lambda(n_2 + c_2)(n + a)/(n + c_0 + c_1 + c_2)(b\lambda + A)$	$(n + a)\lambda/A + b\lambda$
π_2	$n\lambda(2n_0 + 1)/A(2n + 3)$	$n\lambda(2n_1 + 1)/A(2n + 3)$	$n\lambda(2n_2 + 1)/A(2n + 3)$	$n\lambda/A$
π_3	$n\lambda(2n_0 + 1)/2A(n + 1)$	$n\lambda(2n_1 + 1)/2A(n + 1)$	$n\lambda(2n_1 + 2n_2 + 1)(2n_2 + 1)/4A(n + 1)(n_1 + n_2 + 1)$	$n\lambda/A$
π_4	$n\lambda(2n_0 + 2n_2 + 1)(2n_0 + 1)/4A(n + 1)(n_0 + n_2 + 1)$	$n\lambda(2n_1 + 2n_2 + 1)(2n_1 + 1)/4A(n + 1)(n_1 + n_2 + 1)$	$n\lambda(2n_0 + 2n_2 + 1)(2n_2 + 1)/2A(n + 1)(n_0 + n_2 + 1)$	$n\lambda/A$

TABLE 2: MSEs, ALs, and CPs of θ_0, θ_1, θ_2, and θ ($n = 10$).

Method		Para.	θ_0	θ_1	θ_2	θ
MLE		MSE	0.4858	0.8030	0.4865	0.9374
		Boot-AL	2.2414	2.7146	2.2266	1.7920
		Boot-CP	0.9339	0.9294	0.9405	0.9321
Bayes	π_1	MSE	0.4850	0.8012	0.4857	0.9340
		HPD-AL	2.0388	2.5119	2.0425	1.9018
		HPD-CP	0.9663	0.9440	0.9645	0.9369
	π_2	MSE	0.4055	0.5903	0.4061	0.9374
		HPD-AL	1.7980	2.2183	1.8016	1.9034
		HPD-CP	0.9552	0.9399	0.9539	0.9335
	π_3	MSE	0.4678	0.5732	0.3909	0.9374
		HPD-AL	1.8797	2.2193	1.7850	1.9034
		HPD-CP	0.9481	0.9405	0.9569	0.9460
	π_4	MSE	0.3748	0.7042	0.3754	0.9374
		HPD-AL	1.7724	2.3192	1.7760	1.9034
		HPD-CP	0.9527	0.9468	0.9515	0.9405

TABLE 3: MSEs, ALs, and CPs of θ_0, θ_1, θ_2, and θ ($n = 20$).

Method		Para.	θ_0	θ_1	θ_2	θ
MLE		MSE	0.2505	0.4519	0.2523	0.6907
		Boot-AL	1.5795	1.9048	1.5807	1.2957
		Boot-CP	0.9488	0.9483	0.9412	0.9407
Bayes	π_1	MSE	0.2503	0.4512	0.2520	0.6893
		HPD-AL	1.4434	1.7573	1.4382	1.3635
		HPD-CP	0.9832	0.9692	0.9831	0.9415
	π_2	MSE	0.2335	0.3766	0.2350	0.6907
		HPD-AL	1.3512	1.6476	1.3462	1.3640
		HPD-CP	0.9746	0.9447	0.9762	0.9409
	π_3	MSE	0.2551	0.3662	0.2293	0.6907
		HPD-AL	1.3834	1.6486	1.3398	1.3640
		HPD-CP	0.9668	0.9506	0.9777	0.9598
	π_4	MSE	0.2201	0.4260	0.2216	0.6907
		HPD-AL	1.3399	1.6868	1.3348	1.3640
		HPD-CP	0.9614	0.9525	0.9613	0.9498

TABLE 4: MSEs, ALs, and CPs of θ_0, θ_1, θ_2, and θ ($n = 30$).

Method		Para.	θ_0	θ_1	θ_2	θ
MLE		MSE	0.1752	0.3345	0.1771	0.6049
		Boot-AL	1.2849	1.5451	1.2896	1.0510
		Boot-CP	0.9651	0.9516	0.9654	0.9415
Bayes	π_1	MSE	0.1751	0.3341	0.1770	0.6040
		HPD-AL	1.1710	1.4354	1.1727	1.1164
		HPD-CP	0.9919	0.9629	0.9901	0.9427
	π_2	MSE	0.1694	0.2922	0.1712	0.6049
		HPD-AL	1.1197	1.3745	1.1213	1.1167
		HPD-CP	0.9839	0.9723	0.9835	0.9418
	π_3	MSE	0.1814	0.2849	0.1679	0.6049
		HPD-AL	1.1377	1.3750	1.1177	1.1167
		HPD-CP	0.9783	0.9770	0.9851	0.9615
	π_4	MSE	0.1612	0.3228	0.1629	0.6049
		HPD-AL	1.1132	1.3966	1.1148	1.1167
		HPD-CP	0.9896	0.9581	0.9885	0.9638

TABLE 5: MSEs, ALs, and CPs of θ_0, θ_1, θ_2, and θ ($n = 50$).

Method		Para.	θ_0	θ_1	θ_2	θ
MLE		MSE	0.1158	0.2460	0.1161	0.5380
		Boot-AL	0.9947	1.1981	1.0018	0.8227
		Boot-CP	0.9829	0.9578	0.9831	0.9554
Bayes	π_1	MSE	0.1157	0.2458	0.1161	0.5375
		HPD-AL	0.9118	1.1075	0.9071	0.8677
		HPD-CP	0.9955	0.9822	0.9954	0.9724
	π_2	MSE	0.1150	0.2243	0.1154	0.5380
		HPD-AL	0.8874	1.0786	0.8828	0.8679
		HPD-CP	0.9884	0.9900	0.9870	0.9702
	π_3	MSE	0.1209	0.2196	0.1137	0.5380
		HPD-AL	0.8961	1.0791	0.8811	0.8679
		HPD-CP	0.9813	0.9933	0.9901	0.9721
	π_4	MSE	0.1105	0.2417	0.1109	0.5380
		HPD-AL	0.8842	1.0892	0.8796	0.8679
		HPD-CP	0.9938	0.9789	0.9931	0.9717

Table 6: The simulated data when $n = 25$.

(0.0019	1)	(0.0025	1)	(0.0062	1)	(0.0361	0)	(0.0651	2)	(0.0675	1)	(0.1108	2)	(0.1447	2)	(0.1509	1)	(0.1694	2)	(0.1737	0)	(0.1859	1)	(0.1900	2)	(0.2218	0)	(0.2307	2)	(0.2537	2)	(0.2558	2)	(0.2734	2)	(0.2750	0)	(0.3349	1)	(0.3528	1)	(0.3790	0)	(0.3824	2)	(0.3875	1)	(0.5336	1)

TABLE 7: Point estimates and 95% CIs of θ_0, θ_1, θ_2, and θ.

Method	Para.		θ_0	θ_1	θ_2	θ
MLE		MLE	0.8029	1.9270	1.2847	4.0146
		Boot-CI	(0.0408, 1.8099)	(0.1891, 2.8803)	(0.0642, 1.8178)	(0.7671, 4.4112)
Bayes	π_1	Bayes	0.8029	1.9267	1.2845	4.0141
		HPD CI	(0.0898, 1.6991)	(0.3473, 2.9083)	(0.0396, 1.7374)	(0.8141, 4.1876)
	π_2	Bayes	0.8332	1.8937	1.2877	4.0146
		HPD CI	(0.0694, 1.7457)	(0.3070, 2.8296)	(0.0364, 1.7697)	(0.7893, 4.4378)
	π_3	Bayes	0.8492	1.8841	1.2812	4.0146
		HPD CI	(0.0798, 1.7561)	(0.1251, 2.5045)	(0.0315, 1.4392)	(0.6368, 4.1407)
	π_4	Bayes	0.8189	1.9301	1.2656	4.0146
		HPD-CP	(0.0638, 1.4011)	(0.2448, 2.8554)	(0.0456, 1.7418)	(0.8474, 4.3484)

TABLE 8: Point estimates and 95% CIs of θ_0, θ_1, θ_2, and θ.

Method	Para.		θ_0	θ_1	θ_2	θ
MLE		MLE	1.0882e-2	0.4664e-2	1.3214e-2	2.8760e-2
		Boot-CI	(0.7354e-3, 1.2770e-2)	(0.3244e-2, 2.4914e-2)	(0.4077e-3, 1.5990e-2)	(0.7324e-2, 3.5454e-2)
Bayes	π_1	Bayes	1.1882e-2	0.6879e-2	1.3758e-2	3.2520e-2
		HPD CI	(0.2305e-2, 1.1210e-2)	(0.3998e-2, 1.9825e-2)	(0.1861e-2, 1.4366e-2)	(1.0088e-2, 3.8182e-2)
	π_2	Bayes	1.0832e-2	0.4856e-2	1.3073e-2	2.8760e-2
		HPD CI	(0.8317e-3, 1.1676e-2)	(0.2315e-2, 1.8702e-2)	(0.8807e-3, 1.3700e-2)	(0.5246e-2, 3.2002e-2)
	π_3	Bayes	1.0974e-2	0.4817e-2	1.2969e-2	2.8760e-2
		HPD CI	(0.8499e-3, 1.1535e-2)	(0.3160e-2, 1.7152e-2)	(0.8814e-3, 1.4144e-2)	(0.6994e-2, 3.2140e-2)
	π_4	Bayes	1.0803e-2	0.4919e-2	1.3038e-2	2.8760e-2
		HPD-CP	(0.8949e-3, 1.1372e-2)	(0.3365e-2, 1.8866e-2)	(0.6968e-3, 1.4164e-2)	(0.7224e-2, 3.0467e-2)

5. Numerical Simulation and Illustrative Example

5.1. Simulation.

Suppose the common shape parameter λ is known. The initial values for parameters $(\lambda, \theta_0, \theta_1, \theta_2)$ are $(3, 1, 2, 1)$. The initial values for the hyperparameters a, b, c_0, c_1, and c_2 are all 0.001. Take the sample size $n = 10$, 20, 30, and 50. Generate the random samples $(x_1, x_2, \cdots, x_n)$ from $MOGP(\lambda, \theta_0, \theta_1, \theta_2)$ by the following steps:

Step 1: for a fixed value n, generate n samples $u_{01}, u_{02}, \cdots, u_{0n}$ from $GP(\lambda, \theta_0)$, $u_{11}, u_{12}, \cdots, u_{1n}$ from $GP(\lambda, \theta_1)$, and $u_{21}, u_{22}, \cdots, u_{2n}$ from $GP(\lambda, \theta_2)$. Then, we obtain $T_{1l} = \min(u_{0l}, u_{1l})$ and $T_{2l} = \min(u_{0l}, u_{2l})$, $l = 1, 2, \cdots, n$.

Step 2: compute $(x_l, \delta_{0l}, \delta_{1l}, \delta_{2l})$, $l = 1, 2, \cdots, n$, where $x_l = \min(T_{1l}, T_{2l})$, $\delta_{0l} = I(T_{1l} = T_{2l})$, $\delta_{1l} = I(T_{1l} < T_{2l})$, and $\delta_{2l} = I(T_{1l} > T_{2l})$.

Repeat the procedures 10,000 times; we get the values of the mean squared errors (MSEs) of the MLEs, the average lengths (ALs), and coverage probabilities (CPs) of the 95% Boot-P CIs, and the MSEs of the Bayesian estimates, the ALs, and CPs of the 95% HPD CIs, which are shown in Table 2–5. From the results in Table 2–5, we can make the following conclusions.

The MSEs of MLEs and Bayesian estimates decrease as the sample size increases. For given sample size n, the Bayesian estimates based on π_1, π_2, and π_4 are smaller than the MSEs of MLEs. The MSEs of Bayesian estimates of

θ_0 and θ_2 based on π_4 are smaller than that based on π_1, π_2, and π_3. The MSEs of Bayesian estimates of θ_1 based on π_3 are smaller than that based on π_1, π_2, and π_4. The MSEs of Bayesian estimates of θ based on π_1 are smaller than that based on π_2, π_3, and π_4.

The CPs of Boot-P and HPD CIs are all close to 0.95. The ALs of Boot-P and HPD CIs decrease; the associated CPs increase when the sample size increases. The CPs of HPD CIs based on Bayesian estimates are larger than the CPs of Boot-P CIs based on MLEs.

5.2. Illustrative Analysis

5.2.1. Simulated Data.

For illustrative purposes, with initial value for parameters $(\lambda, \theta_0, \theta_1, \theta_2)$ as $(3, 1, 2, 1)$, we use the procedures mentioned above to generate U_0, U_1, and U_2 from $GP(3, 1)$, $GP(3, 2)$, and $GP(3, 1)$, respectively. We then get $T_1 = \min(U_0, U_1)$ and $T_2 = \min(U_0, U_2)$; the latent lifetime of the system is $\min(T_1, T_2)$. The simulated data are listed in Table 6. The MLEs, Bayesian estimates, and associated 95% CIs for parameters θ_0, θ_1, θ_2, and θ are shown in Table 7. From Table 7, all the MLEs and Bayesian estimates of $(\theta_0, \theta_1, \theta_2, \theta)$ are close to the true value.

5.2.2. Real Data.

Use the procedures mentioned above to a real dataset. Kundu and Gupta [13] analyzed the football data of UEFA Champions' League data which are presented in Table 1. From the data, T_1 and T_2 can be regarded as two

dependent failure modes, and $n_0 = 7$, $n_1 = 17$, and $n_2 = 13$. This data have been fitted by Marshall–Olkin bivariate Gompertz distribution (see Wang et al. [23]).

The MLEs, Bayesian estimates, and associated 95% CIs for parameters θ_0, θ_1, θ_2, and θ are shown in Table 8. From Tables 7 and 8, Bayesian estimates under different priors are close to MLEs, and the lengths of 95% Boot-p CIs associated to MLEs are longer than the lengths of 95% HPD CIs associated to Bayesian estimates.

6. Conclusion

This paper discussed the point estimates and CIs for the parameters of the dependent competing risks' model from MOGP distribution. We studied the appropriateness of the posteriors based on conjugate prior and Jeffreys and Reference priors, obtained the Bayesian estimates in closed forms, and constructed the associated HPD CIs. From the simulations results, the use of the Bayesian method can be recommended if the priors are available. The results of the illustrative analysis show that the proposed methods work well; from the lengths of CIs, we can conclude the Bayesian estimates are better than MLEs in general.

Data Availability

The data used to support the findings of the study are available within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] L. Wang, Y. M. Tripathi, C. Lodhi, and X. Zuo, "Inference for constant-stress Weibull competing risks model under generalized progressive hybrid censoring," *Mathematics and Computers in Simulation*, vol. 192, pp. 70–83, 2022.

[2] J. Ren and W. Gui, "Statistical analysis of adaptive type-II progressively censored competing risks for Weibull models," *Applied Mathematical Modelling*, vol. 98, pp. 323–342, 2021.

[3] X. Qin and W. Gui, "Statistical inference of Burr-XII distribution under progressive Type-II censored competing risks data with binomial removals," *Journal of Computational and Applied Mathematics*, vol. 378, Article ID 112922, 2020.

[4] Q. Guan, Y. Tang, and A. Xu, "Objective Bayesian analysis for competing risks model with Wiener degradation phenomena and catastrophic failures," *Applied Mathematical Modelling*, vol. 74, pp. 422–440, 2019.

[5] X. P. Zhang, J. Z. Shang, X. Chen, C. H. Zhang, and Y. S. Wang, "Statistical inference of accelerated life testing with dependent competing failures based on copula theory," *IEEE Transactions on Reliability*, vol. 63, no. 3, pp. 764–780, 2014.

[6] F. Zhang, Y. Shi, and R. Wang, "Geometry of the q-exponential distribution with dependent competing risks and accelerated life testing," *Physica A: Statistical Mechanics and Its Applications*, vol. 468, pp. 552–565, 2017.

[7] Y. Wang and Z. Yan, "Statistical inference on accelerated life testing with dependent competing failure model under progressively type II censored data based on copula theory," *Quality and Reliability Engineering International*, vol. 37, no. 4, pp. 1396–1408, 2021.

[8] M. Wu, Y. M. Shi, and C. F. Zhang, "Statistical analysis of dependent competing risks model in accelerated life testing under progressively hybrid censoring using copula function," *Communications in Statistics - Simulation and Computation*, vol. 46, no. 5, pp. 4004–4017, 2017.

[9] S. M. S. Lo and R. A. Wilke, "A copula model for dependent competing risks," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 59, no. 2, pp. 359–376, 2010.

[10] G. Fang, R. Pan, and Y. Hong, "Copula-based reliability analysis of degrading systems with dependent failures," *Reliability Engineering & System Safety*, vol. 193, Article ID 106618, 2020.

[11] A. W. Marshall and I. Olkin, "A multivariate exponential distribution," *Journal of the American Statistical Association*, vol. 62, no. 317, pp. 30–44, 1967.

[12] Y. Li, J. Sun, and S. Song, "Statistical analysis of bivariate failure time data with Marshall-Olkin Weibull models," *Computational Statistics & Data Analysis*, vol. 56, no. 6, pp. 2041–2050, 2012.

[13] D. Kundu and A. K. Gupta, "Bayes estimation for the Marshall-Olkin bivariate Weibull distribution," *Computational Statistics & Data Analysis*, vol. 57, no. 1, pp. 271–281, 2013.

[14] X. Bai, Y. Shi, H. K. T. Ng, and Y. Liu, "Inference of accelerated dependent competing risks model for Marshall-Olkin bivariate Weibull distribution with nonconstant parameters," *Journal of Computational and Applied Mathematics*, vol. 366, Article ID 112398, 2020.

[15] Q. Guan, Y. Tang, and A. Xu, "Objective Bayesian analysis for bivariate Marshall-Olkin exponential distribution," *Computational Statistics & Data Analysis*, vol. 64, pp. 299–313, 2013.

[16] A. A. Ismail, "Bayes estimation of Gompertz distribution parameters and acceleration factor under partially accelerated life tests with Type-I censoring," *Journal of Statistical Computation and Simulation*, vol. 80, no. 11, pp. 1253–1264, 2010.

[17] M. E. Ghitany, F. Alqallaf, and N. Balakrishnan, "On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively Type-II censored samples," *Journal of Statistical Computation and Simulation*, vol. 84, no. 8, pp. 1803–1812, 2014.

[18] J. M. Bernardo, "Reference posterior distributions for bayesian inference," *Journal of the Royal Statistical Society: Series B*, vol. 41, no. 2, pp. 113–128, 1979.

[19] J. O. Berger and J. M. Bernardo, "On the development of reference priors (with discussion)," in *Bayesian Analysis IV*, J. M. Bernardo, Ed., Oxford University Press, Oxford, England, 1992.

[20] H. Jeffreys, *Theory of Probability*, Oxford Univ. Press, Oxford, 1961.

[21] G. S. Datta and M. Ghosh, "Some remarks on noninformative priors," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1357–1363, 1995.

[22] M.-H. Chen and Q.-M. Shao, "Monte Carlo estimation of Bayesian credible and HPD intervals," *Journal of Computational & Graphical Statistics*, vol. 8, no. 1, pp. 69–92, 1999.

[23] L. Wang, Y. M. Tripathi, S. Dey, and Y. Shi, "Inference for dependence competing risks with partially observed failure causes from bivariate Gompertz distribution under generalized progressive hybrid censoring," *Quality and Reliability Engineering International*, vol. 37, no. 3, pp. 1150–1172, 2021.

Research Article
Connectivity Index of Generalized Uncertain Graph

Jinchan Wang [ID],[1] **Xiulian Gao,**[1] **Xiaoshuang Zhou,**[1] **Changyou Guo,**[2] **and Xiuling Yin**[1]

[1]*School of Mathematics and Big Data, Dezhou University, Dezhou 253023, China*
[2]*School of Information Management, Dezhou University, Dezhou 253023, China*

Correspondence should be addressed to Jinchan Wang; wangjinchan@dzu.edu.cn

In the application of classical graph theory, there always are various indeterministic factors. This study studies the indeterministic factors in the connected graph by employing the uncertainty theory. First, this study puts forward two concepts: generalized uncertain graph and its connectivity index. Second, it presents a new algorithm to compute the connectivity index of an uncertain graph and generalized uncertain graph and verify this algorithm with typical examples. In addition, it proposes the definition and algorithm of α-connectivity index of generalized uncertain graph and verifies the stability and efficiency of this new algorithm by employing numerical experiments.

1. Introduction

In graph theory, the connectivity graph is a fundamental concept. Many phenomena in the real life can be conveniently described as the graphical problems, which are formed by the "points" and the "lines." For example, a roadmap can be interpreted as a graph, the vertices are the junctions, and an edge is the stretch of road from one junction to the next [1]. Graph theory is an important branch of operation research, and it is a young but rapidly maturing subject. Graph theory originated in the eighteenth century. Euler solved the first real problem (seven bridges of Königsberg) using graph theory in 1736 [2, 3], which thereby made him a founder of graph theory.

The information of edges and vertices is completely deterministic in traditional graph theory. However, in applications, due to the complexity of the system, uncertain factors in the graph may appear and produce a new situation. For example, the optimal objective cannot be easily formulated if the weight of the edges is uncertain in some shortest path problems [4], and the classical algorithms fail to solve shortest path problems. Whether two vertices are connected by an edge cannot be accurately determined in real life, which results in the failure of proving the properties of uncertain graph by the classical method. To cope with the

above situation, the concept of uncertain graph was proposed in 2011 in a reference [5].

Sometimes, whether some vertices exist cannot be completely determined and the information of edges cannot be completely determined. For example, in cloud computing networks, the vertices could represent data centers, with edges representing that the two data centers participate in the calculation at the same time. The vertices and the edges in cloud computing networks both have indeterminacy. As a result, it cannot be fully determined whether the graph is connected.

Generally, when we apply probability theory, the exact probability distribution has to be known. However, the probability distribution does not exist in some observational data of small samples. In this case, we must predict the data based on the experience of experts. To properly use empirical data, Professor Liu proposed the uncertainty theory [6] and revised in [7]. In an uncertain environment, the uncertainty theory is a very powerful tool to solve the situations in which only empirical data can be employed.

Based on the work of [8], this study attempts to deal with the situation where the existence of vertices and edges cannot be completely determined. Reference [8] first presented the concepts of generalized uncertain graph and its connectivity index. The study is mainly to compute the

connectivity index of the generalized uncertainty graph. This study puts forward a new algorithm to compute the connectivity index by the uncertainty theory.

There are five sections in the study. Section 1 is the introduction of this study. Section 2 introduces the relevant concepts and relevant properties of the uncertainty theory and graph theory. Section 3 briefly introduces an uncertain graph and puts forward a new algorithm to compute the connectivity index of uncertain graph and then verifies this algorithm with typical examples. Section 4 proposes a new algorithm of connectivity index of generalized uncertain graph based on the concepts of generalized uncertain graph and connectivity index and employs numerical experiments to verify the correctness of this new algorithm. Section 5 is the conclusion of the study.

2. Preliminaries

2.1. Main Theorem in Uncertainty Theory.
Professor Liu Baoding established the uncertainty theory system [9] in 2007, which was revised and improved in 2010 [7]. Today, the uncertainty theory is a very important branch of mathematics.

This study will shortly present the main developments of uncertainty theory in multiple fields. Reference [10] presented the uncertainty process and defined the uncertainty differential equation. Reference [11] proposed uncertain programming. Reference [4] explored the shortest path problem that the weight of the edges is uncertain in 2011. Reference [12] explored the connectivity index of uncertain
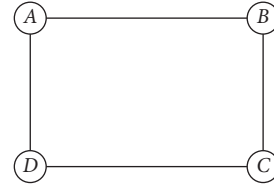
FIGURE 1: A 4-order graph.

graph in 2013. In the next few years, Gao has discussed the cycle index [13, 14], regularity index [8], tree index [15], and α-connectivity index [16] of uncertain graph. Reference [17] discussed some properties of uncertain relations on a finite set in 2014. Reference [18] discussed the distribution function of the diameter in uncertain graph in 2014. There are four main axioms in the uncertainty theory: normality, duality, subadditivity, and product axiom [6, 9]. References [6] and [9, 19] give the concepts and conclusions about the uncertainty theory used in the study.

Theorem 1 (see [19]). *The events $\zeta_1, \zeta_2, \ldots, \zeta_n$ are independent Boolean uncertain variables, i.e.,*

$$\zeta_i = \begin{cases} 1 & \text{if its uncertain measure is } \delta_i, \\ 0 & \text{if its uncertain measure is } 1 - \delta_i, \end{cases} \tag{1}$$

for $i = 1, 2, \ldots, n$, respectively. When g is a Boolean function, $\zeta = g(\zeta_1, \zeta_2, \ldots, \zeta_n)$ is a Boolean uncertain variable such that:

$$M\{\zeta = 1\} = \begin{cases} \sup_{g(\zeta_1, \zeta_2, \ldots, \zeta_n)=1} \min_{1 \le i \le n} v_i(y_i), & \text{if } \sup_{g(\zeta_1, \zeta_2, \ldots, \zeta_n)=1} \min_{1 \le i \le n} v_i(y_i) < 0.5, \\ 1 - \sup_{g(\zeta_1, \zeta_2, \cdots, \zeta_n)=0} \min_{1 \le i \le n} v_i(y_i), & \text{if } \sup_{g(\zeta_1, \zeta_2, \cdots, \zeta_n)=1} \min_{1 \le i \le n} v_i(y_i) \ge 0.5, \end{cases} \tag{2}$$

where y_i is the real number either 0 or 1, and v_i is defined as follows:

$$v_i(y_i) = \begin{cases} \delta_i & \text{if } y_i = 1, \\ 1 - \delta_i & \text{if } y_i = 0, \end{cases} \tag{3}$$

for $i = 1, 2, \ldots, n$, respectively.

2.2. Main Concepts and Terminologies of Classical Graph.
This part will introduce the definition of graph and some main concepts of classical graph theory. At the same time, it will discuss the main theorems about graphs with some examples. The following basic terminology and concepts are from [17].

A graph G means a finite nonempty 2-element set, which consists of a set of vertices and a set of edges. The numbers of vertices of G are called the order, and the numbers of edges of G are called the size. A path is an alternating sequence of vertices and edges if no vertices are repeated. Two vertices are said to be connected if there exists a path between the two

vertices in the graph. A connected graph G means that there exists one path between any two vertices.

Definition 1 (see[17]). *Let G be an n-order and m-size graph, $V(G) = \{v_1, v_2, \ldots, v_n\}$ be its vertices' set, and $E(G) = \{e_1, e_2, \ldots, e_m\}$ be its edges. The symplectic $n \times n$ matrix is the adjacency matrix of G:*

$$A_G = \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix}, \tag{4}$$

where $\delta_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E(G), \\ 0 & \text{if } v_i v_j \notin E(G). \end{cases}$

Example 1. Figure 1 is a 4-order graph.
The adjacency matrix of Figure 1 is as follows:
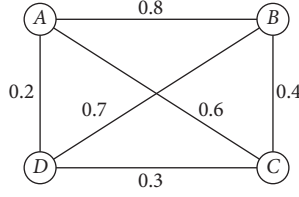
Figure 2: An uncertain 4-order graph.

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}. \tag{5}$$

There is a very important theorem in the classical graph theory, and it is a sufficient and necessary condition to verify whether the graph is a connected graph.

Theorem 2. *If G is a n-order graph and its symplectic adjacency matrix is the A_G, let R be the $n \times n$ matrix, where*

$$R = I + A_G + A_G^2 + A_G^3 + \cdots + A_G^{n-1}. \tag{6}$$

Then, G is a connected graph if and only if $R > 0$.

To analyze graphs and their properties with uncertainty factors, the fundamental concepts are given as follows.

Definition 2. If a symplectic adjacency matrix of the n-order graph is A_G, the connectivity function of G is the following function:

$$C(A_G) = \begin{cases} 1, & \text{if } I + A_G + A_G^2 + A_G^3 + \cdots + A_G^{n-1} > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

3. Uncertainty Graphs

3.1. Basic Concepts. Every vertex and edge of a graph is wholly determined. That is to say, the vertex and the edge either exist or do not exist in the classical graph theory. An n-order graph may be defined as a $n \times n$ symplectic adjacency matrix, and its elements are either 1 or 0. At the same time, the indeterminacy factor will absolutely occur in practical applications. To study this problem, [12] gave the concept of uncertain graph, which means all edges are independent and exist with uncertain measure. That is, the elements are no longer 1 or 0 in adjacency matrix, but a number in [0,1]. For instance, $\delta_{ij} = 0.65$ represents that the uncertain measure is 0.65 about the edge existing between the two vertices, and the uncertain measure is 0.35 about the edge not existing between the two vertices.

Definition 3 (see [12]). The n-order graph G is called an uncertain graph if its symplectic adjacency matrix is as follows:

$$A_G = \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix}, \tag{8}$$

where δ_{ij} means the uncertain measure of the edge between vertexes i and j, for $i, j = 1, 2, \ldots, n$, respectively.

It should be noted that $\delta_{ii} = 0 \, (i = 1, 2, \ldots, n)$ in the uncertain adjacency matrix. Besides, an uncertain adjacency matrix is symmetric if the uncertain graph is undirected; i e., $\delta_{ij} = \delta_{ji}$ for any i and j. Evidently, all the information of an uncertain graph may be contained in the symmetric adjacency matrix.

Example 2. An uncertain 4-order graph is as the following in Figure 2.

The symmetric adjacency matrix of Figure 2 is as follows:

$$\begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix}. \tag{9}$$

We may know that the edge set of G is an uncertain Boolean variable set according to the definition of uncertain graph.

$$E(G) = \left\{\zeta_{12}, \zeta_{13}, \ldots, \zeta_{1n}, \zeta_{23}, \ldots, \zeta_{2n}, \zeta_{34}, \ldots, \zeta_{3n}, \ldots, \zeta_{(n-1)n}\right\}, \tag{10}$$

where $M\{\zeta_{ij} = 1\} = \delta_{ij}$, for any $1 \le i < j \le n$. For the sake of simplicity, we remove these edges ζ_{ij} satisfying $M\{\zeta_{ij} = 1\} = 0$ and denote as $E(G) = \{\zeta_1, \zeta_2, \ldots, \zeta_m\}$.

Definition 4 (see [12]). Let G be an uncertain graph and $E(G) = \{\zeta_1, \zeta_2, \ldots, \zeta_m\}$ be its edge set, and then, connectivity function of G may be denoted as follows:

$$\mathbb{C}(G) = \begin{cases} 1, & \text{if } G \text{ is } a \text{ connected graph,} \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Evidently, $\mathbb{C}(G)$ is a Boolean function. We define connectivity index of an uncertain graph G with $E(G) = \{\zeta_1, \zeta_2, \ldots, \zeta_m\}$ as follows:

$$\partial(G) = M\{\mathbb{C}(G) = 1\}. \tag{12}$$

In other words, the connectivity index is the uncertain measure of the connected graph.

The key is how to obtain the connectivity index in a given uncertain graph. A theorem is proposed to deal with this problem.

Theorem 3. *Let G be an n-order uncertain graph with a symmetric adjacency matrix as follows:*

$$A_G = \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix}. \tag{13}$$

The connectivity index of G is the formula (14) if all edges are independent.

$$\partial(G) = \begin{cases} \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}), & \text{if } \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}) < 0.5, \\ 1 - \sup\limits_{\mathbb{C}(Y)=0} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}), & \text{if } \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}) \ge 0.5, \end{cases} \tag{14}$$

where Y is an uncertain symmetric matrix and its diagonal entries are zero, and Y is denoted as follows:

$$Y = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{pmatrix}, \tag{15}$$

and $y_{ij} = y_{ji}$ is equal to either 0 or 1 and $y_{ii} = 0$, and v_{ij} is denoted as follows:

$$v_{ij}(y_{ij}) = \begin{cases} \delta_{ij}, & \text{when } y_{ij} = 1, \\ 1 - \delta_{ij}, & \text{when } y_{ij} = 0, \end{cases} \tag{16}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$, respectively, and

$$\mathbb{C}(Y) = \begin{cases} 1, & \text{if } I + Y + Y^2 + Y^3 + \ldots + Y^{n-1} > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Proof. Since y_{ij} is independent Boolean uncertain variables for any i and j, Y is the Boolean symmetric uncertain matrix. Therefore, we may get that the function $\mathbb{C}(Y)$ is a Boolean function based on Definition 3. Therefore, according to Theorem 1, we have the following:

$$M\{\mathbb{C}(Y) = 1\} = \begin{cases} \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}), & \text{if } \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}) < 0.5, \\ 1 - \sup\limits_{\mathbb{C}(Y)=0} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}), & \text{if } \sup\limits_{\mathbb{C}(Y)=1} \min\limits_{1 \le i < j \le n} v_{ij}(y_{ij}) \ge 0.5, \end{cases} \tag{18}$$

where Y is the symmetric uncertain matrix as follows:

$$Y = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{pmatrix}, \tag{19}$$

and $y_{ij} = y_{ji}$ is equal to either 0 or 1 and $y_{ij} = y_{ji}$, and v_{ij} is defined as follows:

$$v_{ij}(y_{ij}) = \begin{cases} \delta_{ij}, & \text{when } y_{ij} = 1, \\ 1 - \delta_{ij}, & \text{when } y_{ij} = 0, \end{cases} \tag{20}$$

for any i, j. The connectivity index of an uncertain graph is the uncertain measure of connected graph according to Definition 4. Therefore, Theorem 2 has been proved.

However, the connectivity index is difficult to obtain from Theorem 1 as to other more complex graphs. A new theorem to compute the connectivity index will be given. First, some new definitions will be introduced in the following. □

Definition 5. Assume $A = (a_{ij})$ and $B = (b_{ij})$ be all the $m \times n$ matrices. Then, the logic sum $A \oplus B$ is defined as follows:

$$A \oplus B = (c_{ij}), \tag{21}$$

where $c_{ij} = a_{ij} \vee b_{ij} = \max\{a_{ij}, b_{ij}\}$.

Definition 6. Assume $A = (a_{ij})$ be $m \times k$ matrix and $B = (b_{ij})$ be $k \times n$ matrix. Then, the logic product $A \circ B$ is defined as follows:

$$A \circ B = (c_{ij}), \tag{22}$$

where $c_{ij} = \vee_{p=1}^{k} (a_{ip} \wedge b_{pj}) = \max\limits_{1 \le p \le k}\{\min\{a_{ip}, b_{pj}\}\}$.

Definition 7. Assume $A = (a_{ij})$ be the $n \times n$ matrix. Then, the logic power $A^{(k)}$ is defined as follows:

$$A^{(k)} = A^{(k-1)} \circ A, \tag{23}$$

where I be the $n \times n$ unit matrix, $A^{(1)} = A$ and $A^{(0)} = I$.

Definition 8. Let G be an n-order uncertain graph and its symmetric adjacency matrix be as follows:

$$A_G = \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix}. \tag{24}$$

Then, the reachable measure matrix P of uncertain graph G is defined as follows:

$$P = I \oplus A_G \oplus A_G^{(2)} \oplus A_G^{(3)} \oplus \ldots \oplus A_G^{(n-1)}. \tag{25}$$

Theorem 4. *Let G be an n-order uncertain graph and its uncertain symmetric adjacency matrix be A_G, and its reachable measure matrix be P, and then, the connectivity index of G is the smallest number in the reachable measure matrix P.*

Proof. Since connectivity index of an uncertain graph is the uncertain measure of the connected graph based on Definition 4, a connected graph G is that any two vertices are connected in the graph G. That is, there exists at least a path between any two vertices in the graph G. Thus, we only need to prove that the element $\delta_{ij}^{(k)}$ in $A_G^{(k)}$ is the uncertain measure that there exists at least a path whose length is k from the vertexes i to j in the graph G. In a n-order graph, the

length of a path at most is $n-1$. Then, we can know that Theorem 4 is the right according to the previous theorem and definition.

Now, we begin to prove the conclusion that the element $\delta_{ij}^{(k)}$ in $A_G^{(k)}$ is the uncertain measure that there exists at least a path whose length is k from the vertexes i to j in the graph G using mathematical induction.

Obviously, the conclusion is correct when $k = 0$ or 1.

If $k = 2$, based on Definition 6 and Definition 7, we have the following:

$$\delta_{ij}^{(2)} = \bigvee_{l=1}^{n} \left(\delta_{il} \wedge \delta_{lj} \right), \tag{26}$$

where δ_{il} represents that the edge e_{il} between vertexes i and l exists with uncertain measure δ_{il}, and δ_{lj} represents that the edge e_{lj} between vertexes l and j exists with uncertain measure δ_{lj}, and $\delta_{il} \wedge \delta_{lj}$ represents the uncertain measure that the two edges e_{il} and e_{lj} exist at the same time; i.e., $\delta_{il} \wedge \delta_{lj}$ represents the uncertain measure of the path whose length is 2 from the vertexes i to j via vertex l. So, $\bigvee_{l=1}^{n} (\delta_{il} \wedge \delta_{lj})$ represents the uncertain measure that there exists at least a path whose length is 2 from the vertexes i to j.

According to the above discussion, the conclusion is correct when $k = 2$.

Assume the conclusion is correct when $k = p$, and then, when $k = p + 1$, we have the following:

$$\delta_{ij}^{(p+1)} = \bigvee_{l=1}^{n} \left(\delta_{il}^{(p)} \wedge \delta_{lj} \right), \tag{27}$$

where $\delta_{il}^{(p)}$ represents the uncertain measure that there exists at least a path whose length is p from the vertexes i to l in the graph G, and δ_{lj} represents that the edge e_{lj} between vertexes l and j exists with uncertain measure δ_{lj}, and $\delta_{il}^{(p)} \wedge \delta_{lj}$ represents the uncertain measure that the path whose length is p from the vertexes i to l and edge e_{lj} exists at same time; i.e., $\delta_{il}^{(p)} \wedge \delta_{lj}$ represents the uncertain measure of the path whose length is $p + 1$ from the vertexes i to j via vertex l. So, $\bigvee_{l=1}^{n} (\delta_{il}^{(p)} \wedge \delta_{lj})$ represents the uncertain measure that there exists at least a path whose length is $p + 1$ from the vertexes i to j. Thus, the conclusion is correct when $k = p + 1$.

From what has been discussed above, the conclusion is correct, for $k = 0, 1, 2, \ldots, n-1$, respectively. Therefore, Theorem 3 has been proved. □

Example 3. Figure 2 illustrates the uncertain 4-order graph, and its symmetric adjacency matrix is as follows:

$$A_G = \begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix}. \tag{28}$$

Now, we compute its connectivity index according to Theorem 3.

$$A_G^{(1)} = A_G,$$

$$A_G^{(2)} = \begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix} \circ \begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8 & 0.4 & 0.4 & 0.7 \\ 0.4 & 0.8 & 0.6 & 0.3 \\ 0.4 & 0.6 & 0.6 & 0.4 \\ 0.7 & 0.3 & 0.4 & 0.7 \end{pmatrix},$$

$$A_G^{(3)} = \begin{pmatrix} 0.8 & 0.4 & 0.4 & 0.7 \\ 0.4 & 0.8 & 0.6 & 0.3 \\ 0.4 & 0.6 & 0.6 & 0.4 \\ 0.7 & 0.3 & 0.4 & 0.7 \end{pmatrix} \circ \begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.4 & 0.8 & 0.6 & 0.4 \\ 0.8 & 0.4 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0.4 & 0.6 \\ 0.4 & 0.7 & 0.6 & 0.3 \end{pmatrix}, \tag{29}$$

$$P = I \oplus A_G^{(1)} \oplus A_G^{(2)} \oplus A_G^{(3)},$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 0 & 0.8 & 0.6 & 0.2 \\ 0.8 & 0 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0 & 0.3 \\ 0.2 & 0.7 & 0.3 & 0 \end{pmatrix}$$

$$\oplus \begin{pmatrix} 0.8 & 0.4 & 0.4 & 0.7 \\ 0.4 & 0.8 & 0.6 & 0.3 \\ 0.4 & 0.6 & 0.6 & 0.4 \\ 0.7 & 0.3 & 0.4 & 0.7 \end{pmatrix} \oplus \begin{pmatrix} 0.4 & 0.8 & 0.6 & 0.4 \\ 0.8 & 0.4 & 0.4 & 0.7 \\ 0.6 & 0.4 & 0.4 & 0.6 \\ 0.4 & 0.7 & 0.6 & 0.3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0.8 & 0.6 & 0.7 \\ 0.8 & 1 & 0.6 & 0.7 \\ 0.6 & 0.6 & 1 & 0.6 \\ 0.7 & 0.7 & 0.6 & 1 \end{pmatrix}.$$

So, we can know that the connectivity index of the uncertain graph in Figure 2 is 0.6 according to Theorem 3.

In 2014, the distribution function of the diameter of uncertain graph is discussed in reference [18]. The diameter is an uncertain variable in uncertain graph since the existence of each edge is uncertain. The diameter is a fundamental concept in the graph theory, which means the maximal distance between any two vertices; that is, the diameter in determinate graph $G = (V, E)$ is $\text{diam}(G) = \max_{v_i, v_j \in V} d(v_i, v_j)$. Based on Theorem 3, the following conclusion can calculate the distribution function of the diameter in an uncertain graph.
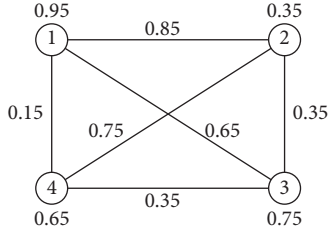
FIGURE 3: A generalized uncertain 4-order graph.

Corollary 1. *Let G be an uncertain n-order graph and its uncertain symmetric adjacency matrix be A_G. Let the matrix D be described as follows:*

$$D = I \oplus A_G \oplus A_G^{(2)} \oplus A_G^{(3)} \oplus \ldots \oplus A_G^{(k)}. \tag{30}$$

Then, the uncertainty distribution of the diameter in the graph *G* is as follows:

$$M\{\text{diam}(G) \le k\} = \min_{1 \le i,j \le n}\{d_{ij}\}, \tag{31}$$

where k ($k \le n-1$) is a positive integer and d_{ij} is the element in *D*.

In the next section, the connectivity of generalized uncertain graph is investigated and analyzed.

4. Generalized Uncertainty Graphs

4.1. Basic Concepts and Algorithms. Not only whether an edge exists between two vertices cannot be completely determined but also whether some vertices exist cannot be completely determined in applications of traditional graph theory. In the study, the indeterministic factors are that whether vertices and edges exist is uncertain. If there are no historical data or experimental data, a random variable cannot be employed to describe this indeterministic factor. Often, we may consult experts and ask them to give a belief degree that the edge exists. The experience data are precisely the research category of uncertainty theory. Therefore, uncertain variable is employed to describe the indeterministic factor in this study.

Definition 9. A n-order graph is called a generalized uncertain graph if its symmetric adjacency matrix is as follows:

$$A_G = \begin{pmatrix} 0 & \delta_{12} & \cdots & \delta_{1n} & \delta_1 \\ \delta_{21} & 0 & \cdots & \delta_{2n} & \delta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & 0 & \delta_n \\ \delta_1 & \delta_2 & \cdots & \delta_n & 0 \end{pmatrix}, \tag{32}$$

where δ_{ij} represents the uncertain measure of edges and δ_k represents the uncertain measure of vertex.

It should be noted that there always are $\delta_{ii} = 0$ in the generalized uncertain adjacency matrix.

From the above, we know that all the information of a generalized uncertain graph is contained in the adjacency matrix.
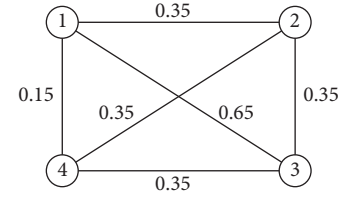
FIGURE 4: Basic graph of Figure 3.

Example 4. Figure 3 is a generalized uncertain 4-order graph as follows.

The adjacency matrix of Figure 3 is as follows:

$$\begin{pmatrix} 0 & 0.85 & 0.65 & 0.15 & 0.95 \\ 0.85 & 0 & 0.45 & 0.75 & 0.35 \\ 0.65 & 0.45 & 0 & 0.35 & 0.75 \\ 0.15 & 0.75 & 0.35 & 0 & 0.65 \\ 0.95 & 0.35 & 0.75 & 0.65 & 0 \end{pmatrix}. \tag{33}$$

Definition 10. Suppose *G* is a generalized uncertain n-order graph, and its adjacency matrix is as follows:

$$A_G = \begin{pmatrix} 0 & \delta_{12} & \cdots & \delta_{1n} & \delta_1 \\ \delta_{21} & 0 & \cdots & \delta_{2n} & \delta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & 0 & \delta_n \\ \delta_1 & \delta_2 & \cdots & \delta_n & 0 \end{pmatrix}, \tag{34}$$

where δ_{ij} represents the uncertain measure of edges between two vertexes and δ_k represents the uncertain measure of vertex. The basic graph $\overline{G}$ of generalized uncertain graph *G* is defined as an uncertain graph with the following adjacency matrix:

$$B_G = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{pmatrix}, \tag{35}$$

where $\lambda_{ij} = \min\{\delta_{ij}, \delta_i, \delta_j\}$.

Example 5. Figure 4 is the basic graph of Figure 3.

The symmetric adjacency matrix of Figure 4 is as follows:

$$\begin{pmatrix} 0 & 0.35 & 0.65 & 0.15 \\ 0.35 & 0 & 0.35 & 0.35 \\ 0.65 & 0.35 & 0 & 0.35 \\ 0.15 & 0.35 & 0.35 & 0 \end{pmatrix}. \tag{36}$$

Definition 11. Suppose that *G* is a generalized uncertain graph with a symmetric adjacency matrix as follows:

$$A_G = \begin{pmatrix} 0 & \delta_{12} & \cdots & \delta_{1n} & \delta_1 \\ \delta_{21} & 0 & \cdots & \delta_{2n} & \delta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \delta_{n1} & \delta_{n2} & \cdots & 0 & \delta_n \\ \delta_1 & \delta_2 & \cdots & \delta_n & 0 \end{pmatrix}. \tag{37}$$

The connectivity function of G is expressed as follows:

$$\mathbb{C}(G) = \begin{cases} 1, & \text{if graph } G \text{ is connected}, \\ 0, & \text{otherwise}. \end{cases} \tag{38}$$

From the above, we can know that $\mathbb{C}(G)$ is a Boolean function. Thus, in a generalized uncertain graph G the connectivity index may be denoted as follows:

$$\partial(G) = M\{\mathbb{C}(G) = 1\}. \tag{39}$$

Namely, the connectivity index is the uncertain measure to which the generalized uncertain graph is connected.

Obviously, the connectivity index of a generalized uncertain graph is exactly equal to the connectivity index of its basic graph.

Based on some of the ideas that we have discussed above, we may know the connectivity index of a generalized uncertain graph, and then, we may compute the connectivity index of a generalized uncertain graph as follows.

Step 1. To write out the adjacency matrix of the generalized uncertain graph G.

Step 2. To compute the adjacency matrix of the basic graph $\overline{G}$ of the generalized uncertain graph G.

Step 3. To compute the connectivity index of the basic graph $\overline{G}$ of the generalized uncertain graph G according to Theorem 3, and thus, the connectivity index of generalized uncertain graph is obtained.

4.2. Examples. In this part, two examples are given to compute the connectivity index of a generalized uncertain graph.

Example 6. Figure 5 is a 2-order generalized uncertain graph G and the uncertain measure of edge is a, and the uncertain measure of vertices is α and β respectively.

In Figure 5, the connectivity index of G is the smallest number of α, β, and a.

Example 7. Figure 6 illustrates a 3-order uncertain graph G, and the uncertain measure of edges is 0.8, 0.7, and 0.6, respectively, and uncertain measure of vertices is 0.75, 0.9, and 0.8, respectively.

As Figure 6 illustrates a 3-order generalized uncertain graph (G), its basic graph is shown in Figure 7. In Figure 6, the connectivity index of (G) is 0.7 by Theorem 1.
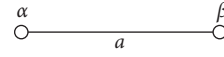
FIGURE 5: A 2-order generalized uncertain graph.
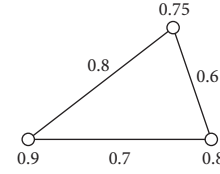
FIGURE 6: A 3-order generalized uncertain graph.
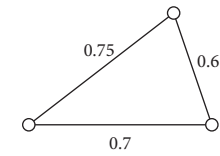
FIGURE 7: Basic graph of the generalized uncertain graph (G) in Figure 6.

Example 8. Figure 3 illustrates a 4-order generalized uncertain graph G, and the uncertain measure of its edges is 0.8, 0.7, 0.6, 0.4, 0.3, and 0.1, respectively, and the uncertain variables of its vertices are 0.9, 0.3, 0.7, and 0.6, respectively.

Because Figure 3 is a 4-order generalized uncertain graph, its basic graph is Figure 4, and the connectivity index of G in Figure 3 is 0.3 by Theorem 3.

4.3. α-Connectivity of Generalized Uncertain Graph

Definition 12. Let G be a generalized uncertain graph with a confidence level α. If the uncertain measure of an edge or vertex in G is not smaller than the confidence level α, the uncertain measure is mapped as 1, or the uncertain measure is replaced by 0. Thus, a sample subgraph of G can be obtained and the sample subgraph is referred to as fundamental subgraph about the confidence level α.

Definition 13. A generalized uncertain graph G is called α-connected if its fundamental subgraph about the confidence level α is connected.

According to the above two definitions, we may easily get an algorithm for how to verify the α-connectivity of a generalized uncertain graph as follows.

Step 1. According to the pre-given confidence level α, write out the fundamental subgraph of the generalized uncertain graph G about the confidence level α.

Step 2. Write down the adjacency matrix $\overline{A_G}$ of the fundamental subgraph of the generalized uncertain graph G about the confidence level α.

Step 3. Calculate the value $\overline{R} = I + \overline{A_G} + \overline{A_G^2} + \overline{A_G^3} + \ldots + \overline{A_G^{k-1}}$, in which k is the order number of the fundamental subgraph of the generalized uncertain graph G about the confidence level α.

Step 4. According to Theorem 1, we can easily know that the generalized uncertain graph G is α-connected if $\overline{R} > 0$; otherwise, it is not α-connected.

Example 9. Figure 8 illustrates a 7-order generalized uncertain graph, and the uncertain measure of its every edge is 0.75, 0.64, 0.69, 0.5, 0.3, 0.2, 0.89, 0.85, 0.55, 0.62, 0.26, 0.9, 0.7, and 0.28, respectively, and the uncertain measure of its any vertex is 0.6, 0.9, 0.65, 0.79, 0.95, 0.8, and 0.86, respectively.

Figure 9 presents the fundamental subgraph of G, and its confidence level is 0.5 when the confidence level α is 0.5. Obviously, it is connected, so the generalized uncertain graph G is 0.5-connected. Figure 10 shows the fundamental subgraph of G, and its confidence level is 0.65 when α is 0.65. Obviously, the graph in Figure 10 is non-connected. Thus, the generalized uncertain graph G is non-connected under the confidence level of 0.65. In the following, we let α be 0.8, we can get a connected fundamental subgraph of G, and its confidence level is 0.8, as shown in Figure 11. Thus, the generalized uncertain graph G is 0.8-connected, but its order is 4.

A major application of α-connectivity of generalized uncertain graph is that the user can determine the size of the graph and the reliability measure according to different requirements.

Definition 14. α-connectivity index of generalized uncertain graph G is defined as follows:

$$\mu(G) = \sup_{\alpha}\{\text{the graph } G \text{ is } \alpha - \text{connected}\}. \tag{40}$$

According to the above discussion, we can easily know the relationship between the α-connectivity index and the connectedness index of a generalized uncertain graph.

Assume G is a generalized uncertain graph, the α-connectivity index of G is equal to its connectivity index.

4.4. Experimental Verification Analysis. We verify the implementation stability and efficiency of the algorithm about the connectivity index of generalized uncertain graph using a large number of experiments. Without a trusted query algorithm according to the reliability index, conditional connectivity query, and so on, the paper only studies the main query methods in different data set scales and the efficiency of different constraint conditions.

The experimental data are implemented in the operating system of Windows 7, the processor of 1.8GHZ, and the memory of 8G condition, using the artificial simulation in the programming environment of Visual C++ 6.0.

In the experiment, we generate randomly some uncertain networks. Table 1 shows the numbers of vertices and edges in these networks. These edges in every uncertain network are randomly created between the two vertices. The reliability of each uncertain edge is randomly created among the real numbers in [0,1], and the reliability of each
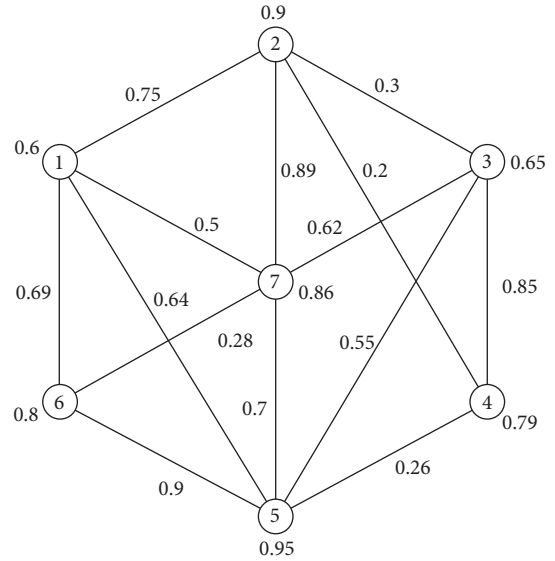
FIGURE 8: Generalized uncertain graph (G) of order 7.
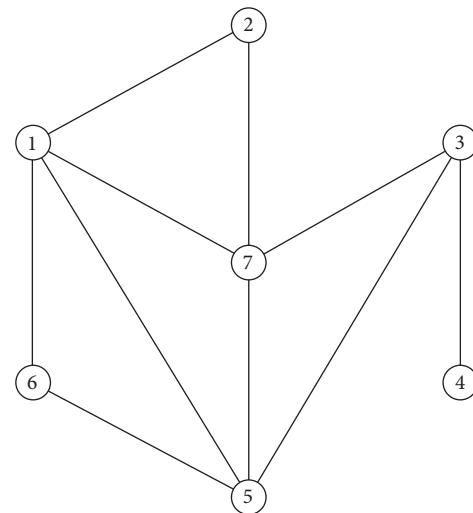
FIGURE 9: An confidence level is 0.5 of fundamental subgraph in Figure 8.

uncertain vertex is randomly created among the real numbers in [0,1].

There are two parts in the experiment. In the following part, we will test the credibility of algorithm of connectivity index generalized uncertain graph if there is no limitation. In Experiment 1, we test the change in algorithm performing required time based on the different scales of edges and vertices. We test the time variation of the algorithm execution under the same number of edges in Experiment 2.

Experiment 1. Figure 12 shows the results of the query results. The execution time of the algorithm has a similar distribution to the change in the number of vertices and edges in the uncertain network.
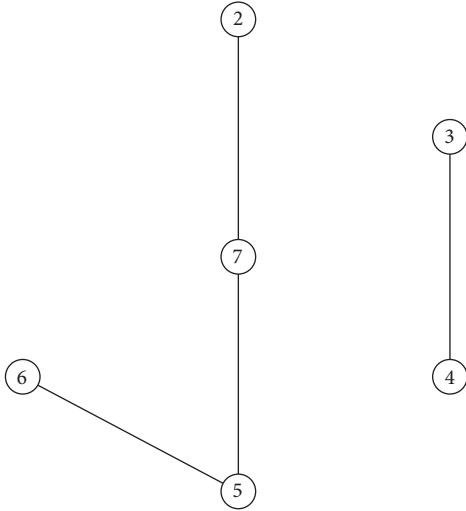
Figure 10: An confidence level is 0.65 of fundamental subgraph in Figure 8.
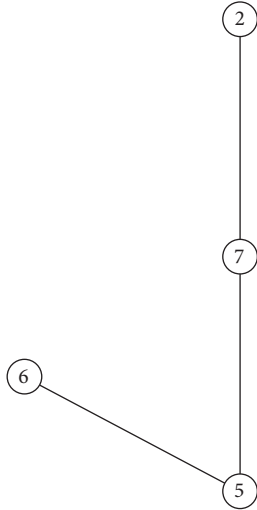
Figure 11: The confidence level is 0.8 of fundamental subgraph in Figure 8.

Table 1: Artificial simulation data of uncertain network.

Name	Vertices	Edges
G1	145	9100
G2	135	7100
G3	125	5100
G4	115	5100
G5	105	4600
G6	95	3600
G7	85	3100
G8	75	2100
G9	65	1800
G10	45	650

Experiment 2. To test the validity of the experiment based on the query algorithm under the user specified the number of edges and vertices in uncertain network. Figure 13 shows
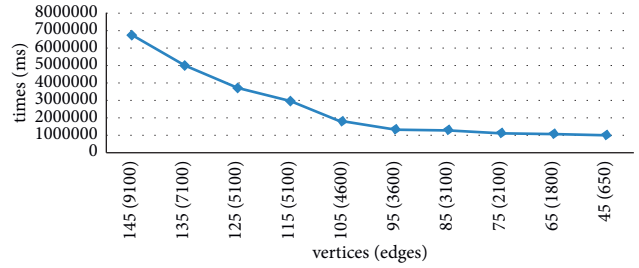
Figure 12: Algorithm execution time versus number of edges and vertices.
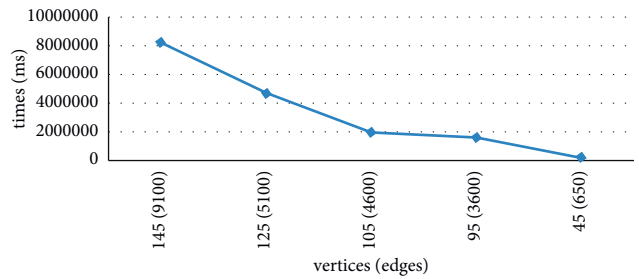
Figure 13: Algorithm execution time versus number of edges and vertices with the reliability constraints ($\alpha = 0.5$).

the query results. Figure 13 shows that it is a similar distribution of the algorithm performing required time, and the change corresponds to a change in the number of edges and vertices in uncertain network.

5. Conclusion

A connected graph is the most basic property in classic graph theory. However, in certain actual applications, different uncertain factors are frequently encountered and must be solved. This study puts forward a definition of generalized uncertain graphs and connectivity index and α-connectivity of generalized uncertain graphs based on the uncertainty theory and proposes an algorithm for connectivity index and α-connectivity by symplectic adjacency matrix of a generalized uncertain graph. This study presents the definition of α-connectivity index of generalized uncertain graphs and obtains the relationship between the α-connectivity index and the connectivity index of a generalized uncertain graph. In the end, the study employs numerical experiments to verify the correctness of this new algorithm.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors state that this article has no conflicts of interest.

References

[1] W. T. Tutte, *Graph Theory*, Cambridge University Press, England, 2001.

[2] L. Euler, "Solutio problematis ad geometriam situs perti-nentis," *Comment. Academiae Sci.I., Petropolitanae*, vol. 8, no. I-7, pp. 128–140, 1736.

[3] L. Euler, "The Königsberg bridges," *Scientific American*, vol. 189, pp. 66–70, 1953.

[4] Y. Gao, "Shortest path problem with uncertain arc lengths," *Computers & Mathematics With Applications*, vol. 62, pp. 2591–2600, 2011.

[5] X. Gao and Y. Gao, "Connectedness Index of Uncertainty Graph," in *Proceedings of the Second International Conference on Uncertainty Theory*, Dayton, OH, USA, Octomber 2011.

[6] B. Liu, *Uncertainty Theory*, Springer-Verlag, Berlin, 2007.

[7] B. Liu, *Uncertainty Theory: A Branch of Mathematics for Modeling Human Uncertainty*, Springer-Verlag, Berlin, 2010.

[8] X. Gao, "Regularity index of uncertain graph," *Journal of Intelligent and Fuzzy Systems*, vol. 27, pp. 1671–1678, 2014.

[9] B. Liu, "Some research problems in uncertainty theory," *Journal of Uncertain Systems*, vol. 3, pp. 3–10, 2009.

[10] B. Liu, "Fuzzy process, hybrid process and uncertain process," *Journal of Uncertain Systems*, vol. 2, pp. 3–16, 2008.

[11] B. Liu, *Theory and Practice of Uncertain Programming*, Springer-Verlag, Berlin, 2009.

[12] X. Gao and Y. Gao, "Connectedness index of uncertainty graph," *International Journal Of Uncertainty, Fuzziness And Knowledge-Based Systems*, vol. 21, pp. 127–137, 2013.

[13] X. Gao, "Cycle index of uncertain graph," *Information: An International Interdisciplinary Journal*, vol. 16, pp. 1131–1138, 2013.

[14] X. Gao, "Cycle index of uncertain graph," in *Proceedings of the Eleventh International Conference on Information and Management Sciences in Celebration of the 60th Birthday of Prof*, Shu-Cherng Fang, Dezhou Shangong, February 2013.

[15] X. Gao, "Tree index of uncertain graphs," *Soft Computing*, vol. 20, pp. 1449–1458, 2016.

[16] X. Gao, C. Guo, X. Yin, and X. Yu, "The computation on α-connectedness index of uncertain graph," *Cluster Computing*, vol. 22, pp. S5691–S5701, 2019.

[17] X. Gao, "Uncertain relations on a finite set and their properties," *Pure and Applied Mathematics Journal*, vol. 3, pp. 13–19, 2014.

[18] Y. Gao, L. Yang, S. Li, and S. Kar, "On distribution function of the diameter in uncertain graph," *Information Sciences*, vol. 296, pp. 61–74, 2015.

[19] B. Liu, *Uncertainty Theory*, Beijing, 2004, http://orsc.edu.cn/liu/ut.pdf.

Research Article

The Application of Digital Technology in the Complex Situation of News Dissemination from the Perspective of New Media Art

Xunxun Jiang [ID]**, Tong Mao** [ID]**, and Jing Tian**

School of Journalism and Communication, Jilin Normal University, Changchun 130021, Jilin, China

Correspondence should be addressed to Xunxun Jiang; jiangxunxun@jlnu.edu.cn

As the continous innovation of new media technology, the media environment of the entire society has undergone profound changes. Digital technology has had a profound impact on the way news is disseminated. It has made a significant impact on the collecting, creation, and distribution of news, as well as the way viewers receive it. As a result, the news media's operation and management style is continually modified. However, in the process of news dissemination, the situations involved are complex and changeable, which leads to different digital technology applications. Aiming at different complex situations in news dissemination under the vision of new media art, this work designs a neural network to optimize the distribution for the required digital technology application schemes. The main work of this paper has the following two points. First, it systematically investigates the current research status of news communication based on digital technology and analyzes the research trends of digital technology and news communication in complex contexts under the vision of new media art. Second, a new neural network is proposed for the optimal application of digital technology for news propagation in different complex situations. This neural network uses an improved particle swarm optimization algorithm and an improved network training strategy to improve the BP network, which can effectively solve the shortcomings of the BP network. A large number of experiments have proved the effectiveness and correctness of this method.

1. Introduction

Digital technology and information science are at the heart of new media. On the basis of the philosophy of mass communication and the creative techniques of modern art, its various domains, including science and art, business and education, culture and the arts, and management, are all integrated through the application of information technology's all-encompassing capabilities. A wide range of media forms are included in new media. New media art's distinguishing features are its communication medium and the digitalization of its disseminating content. Digitization fulfills the creative needs of new media artists by facilitating the collection, access, processing, and sharing of data. It has become a new information carrier in the post-language era, after text and electronic technology. During this period, researchers began to use computer technology for image processing and creation of sound works. At this time, the

forms of digital art showed diversified forms, from static to dynamic interactive works, from creating a virtual world to reflect the initiative of the real world. The development of digital technology makes new media art challenge traditional concepts. Therefore, new media art is a comprehensive media art and performance art that allows the audience to participate [1–5].

News dissemination refers to mass dissemination activities carried out through the news media. It includes not only news gathering, editing, transmission, and reception activities organized by traditional mass media, but also news dissemination activities through the Internet and mobile phones and other digital new media. Books, video games, CD/VCD/DVD, MP3/MP4 music, and other communication methods in the mass media are temporarily not included in the category of news media, because they have nothing to do with the dissemination of real-time news information. Auxiliary news media such as news agencies

and photo agencies will not discuss them temporarily, because they do not directly address a broad audience [6–10].

News communication has always been closely related to development of technology; communication technology is the first driving force for the development of communication, because it determines the renewal of the media, promotes the transformation of communication methods, and leads to the evolution of communication concepts and the development of high-energy journalists. Everywhere we go in the history of news reporting we see the traces of technical progress. Modern electronic technology has made it possible for news dissemination to outpace text delivery in terms of speed and efficiency. This has entered a new field of sound and image dissemination, further expanding the geographical scope and number of audiences of news dissemination. There has been a dramatic shift in the way news is disseminated since the introduction of digital technologies and the subsequent advancement of computer, multimedia, communication, and network technologies. The digitization and networking of news dissemination have changed the existence of news dissemination, the transmission content, and the temporal and spatial relationship of human activities and will create a broader space for news dissemination [11–15].

With the complexity of news dissemination situation, the digital technology required in the dissemination process is also different. For different communication situations, how to design the optimal digital technology application program is an important and novel topic in the process of news communication. With the development for computer science, neural networks can efficiently model this task. Therefore, this article designs a neural network based on computer neural network technology. It can optimize and customize digital technology application schemes for the complex situations of news dissemination under the vision of new media art.

The contributions of this article are as follows: (1) this article applies neural network to the complex situation of news dissemination, which is used to customize the optimal digital technology application scheme in this situation. (2) For the customization of digital technology application schemes, this paper designs an artificial neural network. This network combines the improved particle swarm algorithm with BP network and optimizes the training strategy of BP network accordingly. This can effectively improve the customization performance of digital technology applications.

2. Related Work

Foreign research in this area had been carried out relatively early. Sociologists and technical experts began to study this issue and achieved some results. Literatures [16–19] discussed the application of digital technology and outlined a rough outline for the future information society. Since the development of digital technology was still in its infancy at that time, these works had very little to do with the application of digital technology in news dissemination. It was not until the next few years that the literature [20] clearly put forward the argument that "the world of media has changed." And in the book, the concepts of digital TV, multimedia, virtual reality, and so on were mentioned. Other countries with rapid development of digital technology had also produced a number of research results on emerging online media. It mainly included literatures [21–23] and so on.

Before the mid-1990s, related research in my country was progressing slowly. After the mid-1990s, with the gradual advancement of my country's digitalization process, related foreign academic works had been translated and published successively, such as the related documents mentioned above. Some domestic scholars had begun to get involved in this research field and had achieved great research results. This was mainly manifested in the following: (1) related works and papers continued to emerge. Literature [24] analyzed the content of papers published in Journalism and Communication Research, International Journalism, and Journalism University from 1996 to the end of the 20th century. It found that articles related to media technical analysis showed an overall increasing trend, increasing from the previous single-digit percentage to more than 25%. A large number of relevant research articles had been published on the online platform "China Journalism and Communication Review" jointly established by Sina.com, Zhejiang Online, and the School of Journalism and Communication of Zhejiang University. (2) Among the papers presented at large-scale academic conferences held by the journalism and communication circles in my country, the proportion of papers related to the research on new technologies of journalism and communication had risen sharply. The School of Journalism and Communication of Tsinghua University held a seminar with the theme of "Media Reform in the Digital Background." The central issue discussed in this seminar was the impact of digital technology on media. The participating experts discussed the latest development trends of digital media, the behavior and characteristics of people using new media, and various issues related to this. (3) The National Social Science Fund was also very concerned about this field. Since 1996, related topics had been established almost every year, such as the 1996 project "Multimedia Technology and News Dissemination," the 1998 project "Research Report on the New Development of Network Communication and Its Countermeasures," and "Analysis and Development Forecast of the Digital Status Quo of News Dissemination Means." The 8 projects established in 2000 included "Research on the Influence of the Internet on Information Dissemination and People's Spiritual and Cultural Life" [25, 26]. Among relevant domestic researches, the more influential works included literatures [27–30] and so on.

In the era of digital information, news dissemination tools had achieved a high degree of uniformity, and computers and networks had become important tools for processing and transmitting news information. In the digital age, traditional writing tools and transportation tools were no longer needed for the processing and transmission of information. Computers and networks had completely replaced them. Moreover, the information carrier of news dissemination had also

been completely unified. Any information object, whether it was numbers, words, or symbols, or sounds, graphics, or images, only needed to be carried by bits. At the same time, digital information for news dissemination on the Internet is open, and people's acceptance and consumption of information had also shifted from a passive state in the materialized information age to a more active and interactive state. Therefore, it could be said that digital technology had brought a new stage of news dissemination. At this stage, the essence of news dissemination had not changed, and the function of news dissemination in monitoring the environment, coordinating society, and transmitting cultural heritage had not changed much. But the way and content of news dissemination were constantly changing. Literature [31] believed that the content of one communication tool was often another communication tool. The content of writing was speech, just as handwriting was the content of printing, and printing was the content of telegram. Literature [32] pointed out that when technology used symbolic signs or found a place in a special social structure, it became a medium, a social and intellectual resource created by a machine.

3. Method

This article uses neural networks to specify the best digital technology solutions for different news dissemination scenarios called DTPDN. For example, when the content of news dissemination is sound, and the media is broadcasting, voice signal processing is the most suitable digital technology. When the content of news dissemination is video, and the transmission medium is TV, video processing technology is the most suitable digital technology. In short, this article is to use neural network to model these attribute parameters through the parameter attributes of various complex situations under the vision of new media art and finally output the most optimized digital technology solution. The model used in this article is BP neural network, and it is improved by PSO and training optimization strategy. This can make the neural network more accurate in the formulation of digital technology solutions. The structure is illustrated in Figure 1.

3.1. Improved PSO. The classic BP network model's ability to forecast outcomes is not perfect. There is a reason for this: the BP network algorithm's initial threshold and weight parameters are often arbitrarily set. A local minimum state of the BP network is easy to enter, which results in low model prediction accuracy. This chapter uses an improved particle swarm optimization approach to optimize the initial parameters of the BP neural network model, which successfully increases the model's accuracy.

3.1.1. Particle Swarm Optimization Algorithm. For population assumed by particle swarm optimization, each particle is abstracted as an individual that can move independently and has a certain speed. In the objective function's solution space, each particle is a solution that is constantly

moving. The initial speed and initial direction of the particle flight can be randomly generated as the initial state of each individual. And based on the overall flight experience in the population and the individual's own flight experience to jointly determine the flight plan and status of the next iteration, in the actual problem to be solved by the PSO algorithm, it is necessary to determine the optimized objective function, also called the fitness function. Each individual particle judges and evaluates its own flight search effect according to the value of the fitness function. During the flight search process, the point where each individual encounters the optimal fitness function value is the optimal solution pBest. In the same way, by comparing the individual optimal pBest found by all the individual particles, the global optimal gBest can be compared. In the continuous iterative operation of PSO, each individual particle dynamically adjusts its flight search plan based on individual optimal pBest and global optimal gBest. Such a process is actually similar to the evolutionary process of species. Because each iteration produces a better population, PSO algorithm is also called evolutionary algorithm. The basis of PSO is to use the collective wisdom of the population to conduct a global search and compare local optimal solutions according to the fitness function. These processes do not need to derive the objective function, nor do they need to have too many assumptions and conventions.

The population is composed of some abstract individual particles. Each individual has no mass or volume, and the number of individuals is the dimension of the population. Assume that the dimension for a certain group is n, the position of the i-th particle in the group is a vector $X_i = [x_1, x_2, \ldots, x_n]$, and the flying speed of the particle is a vector $V_i = [v_1, v_2, \ldots, v_n]$. In PSO, each individual particle maintains the optimal value of the individual searched by itself. Then, in the process of individual particle flight search, it is necessary to compare with the global optimal value to judge and update its position. Its essence is to exchange global search result information, which is also the global search capability of particle swarm algorithm:

$$
\begin{aligned}
v(t+1) = v(t) &+ a_1 r(\,)(\text{pbest}(t) - p(t)) \\
&+ a_2 r(\,)(\text{gbest}(t) - p(t)),
\end{aligned} \tag{1}
$$
$$
p(t+1) = p(t) + v(t+1),
$$

where $r(\,)$ is a random function, $v(t)$ is the velocity expression of the particle, and a_1 and a_2 represent the learning speed of particles.

In actual optimization applications, it is often hoped that the particles can have a larger acceleration when the standard particle swarm optimization method is just started, thereby improving the global search ability. When the population obtains certain information, the search space is quickly converged to a smaller range. Then, reduce the particle's movement speed, so that the particle can be searched finely in the local area. Therefore, a nonnegative inertia weight is added to the particle's velocity term to control the speed of the current particle's velocity transformation:
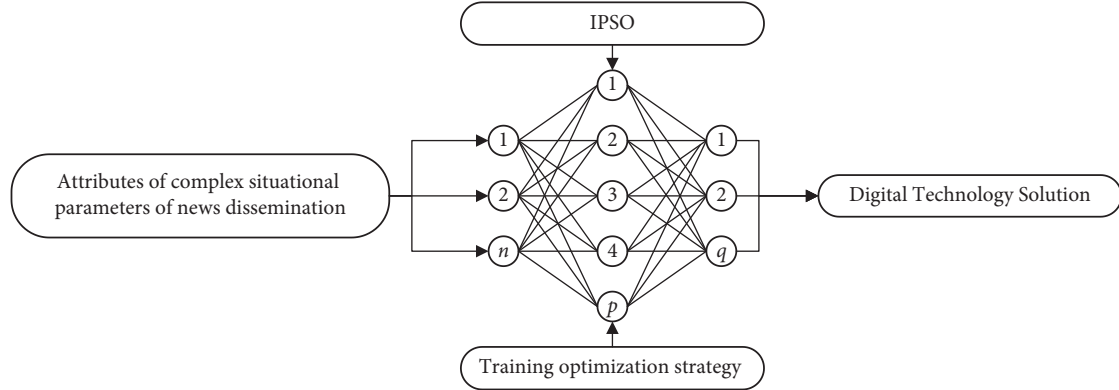
Figure 1: The structure of DTPTN.

$$v(t + 1) = wv(t) + a_1 r(\,)(\text{pbest}(t) - p(t))$$
$$+ a_2 r(\,)(\text{gbest}(t) - p(t)), \tag{2}$$

where w is the inertia weight. The larger the inertia weight, the greater the potential for global optimization, and the lesser the ability for local optimization. The lower the value, the weaker the ability of global optimization, and the better the ability of local optimization. Dynamically changing the inertia weight can obtain better optimization ability than that of the fixed inertia weight. The inertia weight can change linearly during the particle swarm search process, or it can be dynamically changed according to a certain measurement function of the particle swarm performance. At present, the linearly decreasing weight strategy is widely used:

$$w = \frac{(w_s - w_e)(G_k - g)}{G_k} + w_e, \tag{3}$$

where G_k is the maximum number of iterations, w_s is the initial value of the inertia factor, and w_e is value of the inertia factor when the population is iterated to the maximum number of evolutions.

The advantages of particle swarm optimization algorithm are extremely significant. (1) The superiority of the idea of swarm intelligence is that it is not easy to fall into a local optimization, and it has the ability to search and optimize globally. (2) The population converges to the global optimum fast. (3) The search of each particle is parallel, except when exchanging information; the movement between the particles can be regarded as unrelated to each other, and they are all moving on the path they think is correct. (4) There are no special assumptions and restrictions on the input data, which makes the particle swarm optimization algorithm suitable for solving almost all optimization problems. (5) There is almost no need to modify any parameters; you can input into the model. (6) The standard particle swarm algorithm has a simple idea and does not have high requirements on the input data itself. The computational time complexity of PSO is much lower than that of the genetic algorithm.

3.1.2. Improvement Strategy. The standard particle swarm optimization algorithm converges to a local extremum and

cannot escape when it is calculated and solved. Because every time the population exchanges information, the search scope is narrowed. At this time, it is possible that the particle has not searched all the regions, or it is too late to find the optimal solution. The search area is narrowed down by the population, resulting in the optimal solution being excluded from the search range, and it can only converge to a local extremum. Many documents have studied this problem, but the nature of the problem has not changed, and the scope of the global search has been narrowed down. Therefore, this article attempts to improve this problem. When each particle searches for its own extreme value, it is compared with the extreme value found by population. If the extremum searched by the population is better than the extremum searched by the particle itself, the standard particle swarm method strategy is to update position and coordinate vector of the particle to the position and coordinates of the extreme point found by the population. At this time, the search range of the population will be narrowed down to the vicinity of the extreme point. However, the strategy adopted in this paper is to add random variation factors after the population finds the extreme points. After the particle position and coordinates are updated, the global search range is still maintained. It avoids the global optimal solution being excluded from the search area. At the same time, modify the adaptive function and modify the conditions for the particles to end the search, so that the particles can finally exit the iteration. The improved PSO in this paper can expand the search range of particles in the solution space, so that the particles have a certain random distribution in the global search range. This makes the algorithm converge to the optimal solution faster, while also ensuring that it will not converge locally.

To overcome the premature convergence, it is easy to fall local optimum. After the optimal information is exchanged, a dynamic position change is added when narrowing the search range. This allows the particles to have different random search capabilities at different stages. At the same time, some checking mechanisms are added to determine whether the particles have a tendency to locally converge. This can adjust the movement speed and position of the particles in time, so that the entire particle swarm can continue to maintain a faster convergence rate. This will not

sacrifice the global optimization capability of PSO itself. The specific design is as follows.

The variation factor of the particle is a binary random function with respect to the number of iterations and time. Use the number of iterations to consider, and add time as an independent variable. When the population is initialized, time t is 0 at this time. And within a certain period of time, no mutation occurs. Therefore, the significance of the independent variable of time is to control the timing of mutation. At the same time, the consideration of the number of iterations is that if iterations reached a certain value, the population still does not converge, and then mutation is turned on. This allows the particles to have a global search range.

But the fitness converges to a certain value, and the number of times of convergence reaches a certain number of times. At this time, it may be converged to the local optimum, and mutation is turned on at this time. This can help particles escape from local extreme points.

When the position of the particle is always in a certain area, and it is difficult to escape, it is considered that the population has the possibility of converging to the local optimum, and mutation can be turned on. The specific mutation factor operation is to modify the position of the particles with formula of standard PSO:

$$p(t + 1) = P \times r() \times p(t), \tag{4}$$

where P is the core probability factor, which determines the intensity of variation in the next iteration of the particle.

According to the improved ideas and design of PSO, the specific implementation steps of improved PSO (IPSO) are as follows: (1) initialize particle swarm and initialize dimension of the population. The initial velocity, initial position, and velocity extremes of individual particles are initialized to the optimal value of each individual particle. Set the fitness function and the maximum iterations to ensure that the algorithm can finally end. (2) According to global optimal value and individual optimal value, update the speed and position of each individual in the population, and judge whether the speed of each particle is out of bounds. If the boundary is exceeded, the velocity threshold is used to limit the velocity of the particles. (3) Update the global optimal value of the population based on the individual optimal value of all individuals in the population. (4) According to the differences in speed and location of the population before and after evolution, if the population has the possibility of converging to a local optimum, compare the iterations with the number for convergences to the local. According to the dynamic variation formula, the position of the particles is adjusted dynamically and has a certain uniform distribution in the whole world. (5) If the end condition is not reached, then go to step 2. (6) The algorithm operation ends. The final judgment condition is that the fitness of the particles meets the expected error range or the algorithm encounters an abnormality. This adaptive mutation will be called when the position update and speed update of the particles in the population are used to increase diversity of population and also maintain a better convergence speed.

3.2. Training Optimization Strategy.

The training of BP uses error backpropagation method. The change direction of the weight threshold depends on the change direction of the error function. This method begins to decline rapidly, but when the error function enters a flatter curve, it changes slowly and falls into a local extreme. Aiming at this defect, this paper proposes an improved BP method and designs an improved additional momentum algorithm and an improved adaptive learning rate adjustment method.

3.2.1. Improved Additional Momentum Algorithm.

The standard BP algorithm only relies on a simple static optimization method to modify the weight and does not consider the gradient direction at the previous time. As a result, the network converges slowly and oscillates during the learning process. Literature [33] proposed a BP learning algorithm to increase the momentum term. The weight change in this method is obtained by summing the appropriate ratio of the current negative gradient change of the error surface and the weight change used in the previous iteration correction. The weight adjustment formula of the additional momentum term is

$$\Delta w_{ij}(n + 1) = -(1 - \alpha)\beta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(n),$$

$$\Delta b_{ij}(n + 1) = -(1 - \alpha)\beta \frac{\partial E}{\partial b_{ij}} + \alpha \Delta b_{ij}(n), \tag{5}$$

where α is the momentum factor.

In order to increase learning speed, the momentum term can store the value of a change in the weight of the link at the previous time. Additionally, the added momentum component acts as a buffer and smoother function. During training, the inertial effect can be used to reduce the occurrence of oscillations. It is also advantageous for escaping the flat terrain because of the additional momentum period. If the training has entered the flat area of error surface, then $w(n + 1) = \Delta w(n)$, thus preventing the situation of $\Delta w(n + 1) = 0$, making the model jump out of the local minimum. To put it simply, the momentum component added in the additional momentum term technique is essentially a damping term that can help buffer and smooth out the learning process, minimizing the oscillation trend and therefore enhancing the network's converged performance. The addition of a momentum term can effectively reduce the network's susceptibility to local error surface features. It can also be effectively controlled when the network falls into a local minimum. However, when the error surface becomes steeper, it becomes difficult to select the learning rate. To solve this problem, many works have conducted research and proposed some solutions, for example, the adaptive learning rate adjustment method. Although this method improves the difficulty of learning rate selection to a certain extent, it still has the problem of slow learning rate update. In order to update the additional momentum in time, the judgment conditions of the improved use weight correction formula designed in this paper are

$$\alpha = \begin{cases} 0, & E(n+1)^3 > 1.04E(n)^3, \\ 0.95, & E(n+1)^3 < E(n)^3, \\ \alpha, & \text{others.} \end{cases} \quad (6)$$

Judging by this condition, the value of additional momentum is updated. To follow up on the change of the error in time, adjust the direction of the change of the weight threshold.

3.2.2. Improved Adaptive Learning Rate. It is impossible to change the conventional BP algorithm's learning rate after it has been established. BP neural network learning process is slowed down by this issue. The selection of the learning rate should not be too wide-ranging. The network may not be able to converge if the learning rate is set too high and oscillates throughout the training period. In other words, if the rate of learning is too low, the convergence speed cannot be guaranteed. It is difficult to make major modifications in a reasonably flat error surface, because some sections of the BP network error surface are uneven. Even if the adjustment amount of the weight is increased, it is difficult to reduce the error quickly. In this case, a larger learning rate is required. The gradient of the error changes very rapidly in a relatively steep area, and a smaller learning rate needs to be selected at this time.

This paper proposes an improved method of adaptively adjusting the learning rate. As a result, the normal BP algorithm's convergence characteristics are improved by this algorithm's ability to adaptively adjust its learning rate to manage the BP neural network's gradient descent speed in learning. It is vital to maintain a large learning rate in order to provide stability in the learning process, which is based on particular principles. The learning rate adaptive adjustment formula is

$$\beta(n+1) = \begin{cases} 1.05\beta(n), & E(n+1)^3 < E(n)^3, \\ 0.7\beta(n), & E(n+1)^3 > 1.04E(n)^3, \\ \beta(n), & \text{others.} \end{cases} \quad (7)$$

The learning rate is automatically adjusted based on error. When the error becomes a certain multiple compared with the previous error, the learning rate will become smaller. Otherwise, the learning rate remains unchanged. If the error becomes smaller, increase the learning rate. This method can keep the learning rate at the maximum acceptable value throughout the training process. The learning rate is updated quickly to ensure that the network is always trained at the maximum acceptable learning rate. This method can dynamically process the network, and it can make the learning rate reasonable, so as to speed up convergence speed. This can ensure the stability of BP network and the convergence of BP algorithm.

4. Experiments and Discussion

4.1. Dataset. This article uses a self-built dataset for experimentation. The input of this dataset contains four first-level indicators of news dissemination situation, and each first-level indicator contains several second-level indicators. The specific indicator is illustrated in Table 1.

TABLE 1: Parameter index of news dissemination situation.

First-level	Second-level
News dissemination subject	Government
	Organization/institution
	Individual
News dissemination medium	Radio
	Television
	Internet
News dissemination content	Sound
	Image
	Video
News dissemination audience	Old people
	Middle-aged people
	Young people

This article quantifies these indicators as the input of network, and the output is the corresponding digital technology application program, such as voice processing technology and video processing technology. The self-made dataset in this article contains a total of 16,392 pieces of data, of which 80% is training set and 20% is testing set. The evaluation indexes applied in this work are precision, recall, and F1 score.

4.2. Evaluation on Training Convergence. The convergence of network and performance of the convergence are two critical performance measures in the neural network. Studying training loss and precision is used to determine if this is a contributing element. Results are illustrated in Figure 2.

The training loss lowers, and the training precision increases as the network training advances. Network convergence occurs when the number of iterations exceeds 100. Network reliability and robustness may be ensured by achieving convergent network training with our proposed algorithm.

4.3. Comparison with Additional Methods. In order to determine the success of our deigned method, we compare this model to other methods such as logistic regression (LR), decision tree (DT), random forest (RF), and support vector machine (SVM). Precision, recall, and the F1 score are some of the comparison metrics. The experimental result is illustrated in Table 2.

It is not difficult to see that, compared to other similar methods, our method can obtain the highest performance on the three evaluation indicators. Specifically, 0.928 precision, 0.904 recall, and 0.915 F1 score can be obtained, respectively. Compared with the best method listed in the table, our method can obtain 2.5% precision gain, 2.7% recall gain, and 2.2% F1 score gain. This shows that our method is able to formulate accurate digital technology application plans for various complex communication environments under the new media art perspective. These data illustrate the effectiveness and reliability of our method.
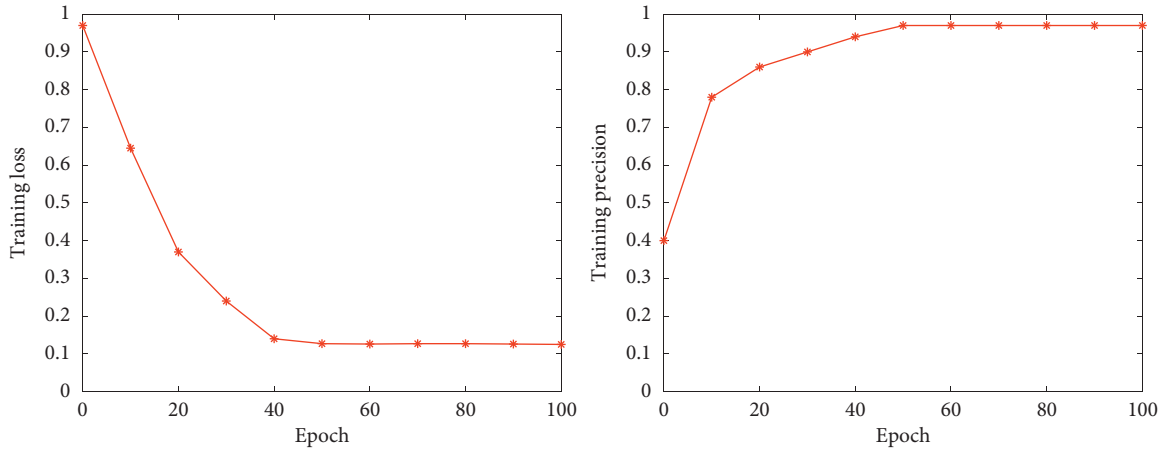
Figure 2: Training loss and training accuracy.

Table 2: Comparison with additional methods.

Method	Precision	Recall	F1
LR	0.827	0.808	8.811
DT	0.852	0.843	0.847
RF	0.886	0.859	0.864
SVM	0.903	0.877	0.893
DTPDN (ours)	0.928	0.904	0.915
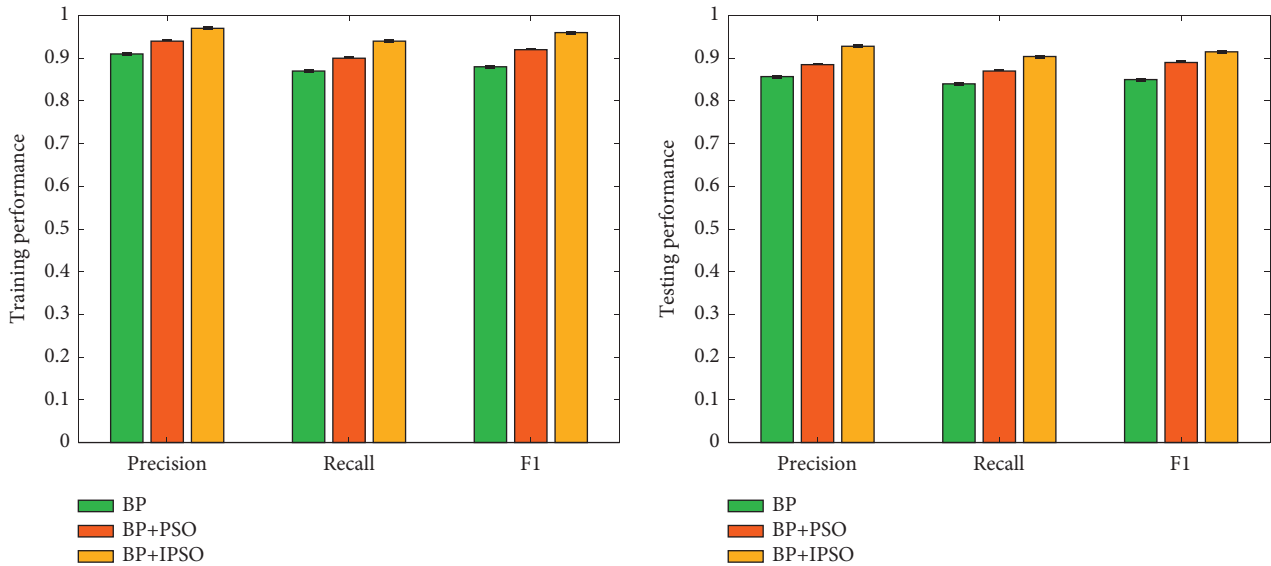
Figure 3: Evaluation on IPSO.

4.4. Evaluation on IPSO. As mentioned earlier, this paper designs an IPSO algorithm for the problem that BP networks tend to fall local optimal solutions to assign values to the initial weights and thresholds for networks. To verify the effectiveness, this work will compare the performance of not using genetic optimization algorithm, using traditional PSO, and using IPSO. The experimental results are shown in Figure 3.

Obviously, whether it is during network training or testing, the performance of the traditional BP network is the lowest. But when the PSO algorithm is used, the performance of the network has been improved to a certain extent, and the improvement is limited. If the PSO algorithm is improved, the IPSO algorithm is constructed. After combining it with the BP network, a further performance improvement can be obtained. Compared with the traditional
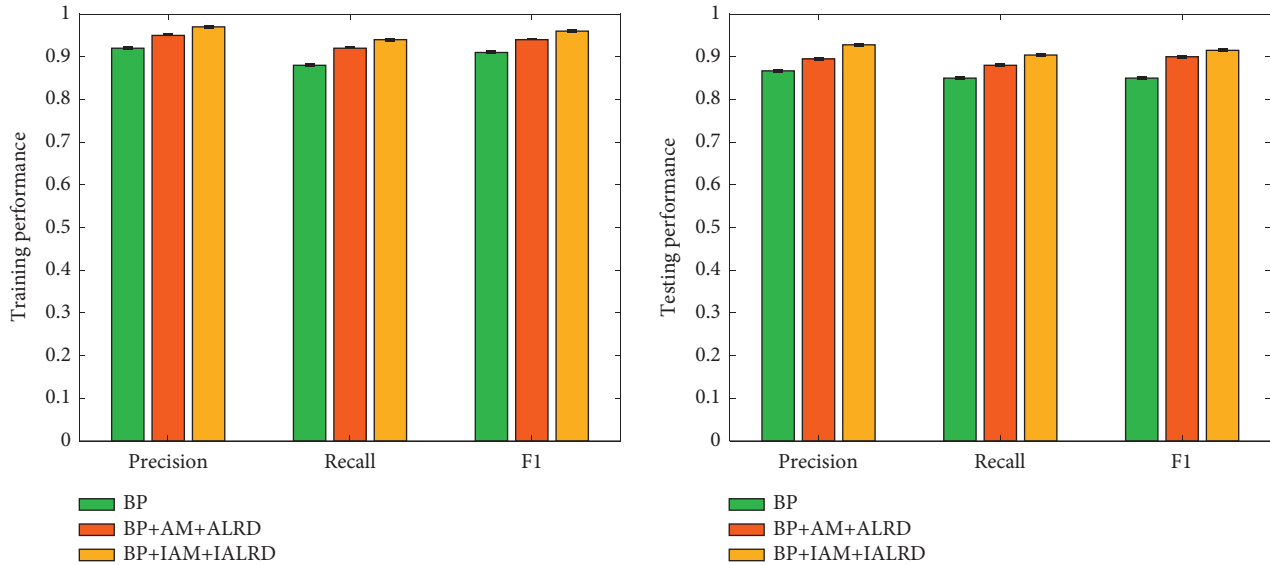
Figure 4: Evaluation on training optimization strategy.

PSO algorithm, the testing performance can be improved by 4.3%, 3.4%, and 2.5%. This proves the correctness and effectiveness of IPSO and can improve the accuracy of the formulation of the optimal digital technology application program in the news dissemination of complex situations.

4.5. Evaluation on Training Optimization Strategy. As mentioned earlier, the training of BP neural network usually adopts the error backpropagation method. The change direction of the weight threshold depends on the change direction of the error function. In this method, the loss drops rapidly when the network is first trained, but when the error function enters a flatter curve, the change is slow. Aiming at this defect, this paper proposes an improved BP algorithm and designs an improved additional momentum (IAM) algorithm and an improved adaptive learning (IALRD) rate adjustment method. To verify the effectiveness of this method, this work will compare the performance of not using training optimization strategy, using traditional additional momentum (AM) algorithm with adaptive learning (ALRD) rate adjustment method, and using improved additional momentum (IAM) algorithm with improved adaptive learning (IALRD) rate adjustment method. The experimental results are illustrated in Figure 4.

Obviously, whether it is during network training or testing, the performance of the traditional BP network is the lowest. But when the training optimization strategy with improvement is used, the performance of the network has been improved. But the improvement is limited. If the training optimization strategy is improved, after combining it with the BP network, a further performance improvement can be obtained. Compared to traditional optimization, the testing performance can be improved by 3.2%, 2.3%, and 2.5%. This proves the correctness and effectiveness of the designed training optimization strategy in this paper and can improve accuracy of formulation of optimal digital

technology application program in the news dissemination of complex situations.

5. Conclusion

With the changes in society and new media art technology, the contemporary media environment has undergone profound changes. Moreover, the development and application of digital technology have brought tremendous changes to news dissemination under the vision of new media art. It has greatly improved every aspect of news dissemination and the way the audience accepts it. This factor keeps the development model of news media updated. In the process of news dissemination, the situations involved are complex and changeable, which directly leads to different applications of digital technology. In view of the different complex situations in news dissemination from the perspective of new media art, this paper proposes an improved artificial neural network to formulate the required digital technology application program. First, this article systematically examines the status quo of digital technology in news dissemination from the perspective of new media art. It also analyzes the research trend of digital technology and news communication in the complex context under the new media art vision. Secondly, this article proposes a new type of neural network to optimize the application of digital technology in news dissemination in different complex situations. The neural network uses an improved particle swarm optimization algorithm and an improved network training strategy to improve the BP network, which can effectively solve the shortcomings of the BP network. Massive experiments have proved the effectiveness and correctness of this method.

Data Availability

The datasets used are available from the corresponding author upon reasonable request.

Conflicts of Interest

Acknowledgments

References

[1] J. M. Twenge, T. E. Joiner, M. L. Rogers, and G. N. Martin, "Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time," *Clinical Psychological Science*, vol. 6, no. 1, pp. 3–17, 2018.

[2] S. Edgerly, E. K. Vraga, L. Bode, K. Thorson, and E. Thorson, "New media, new relationship to participation? a closer look at youth news repertoires and political participation," *Journalism & Mass Communication Quarterly*, vol. 95, no. 1, pp. 192–212, 2018.

[3] J. McMullan, "A new understanding of "New Media": online platforms as digital mediums," *Convergence: The International Journal of Research into New Media Technologies*, vol. 26, no. 2, pp. 287–301, 2020.

[4] S. Boccalini, P. Bonanni, F. Chiesi et al., "The experience of VaccinarSinToscana website and the role of new media in promoting vaccination," *Vaccines*, vol. 8, no. 4, p. 644, 2020.

[5] R. Ritonga and I. Syahputra, "Citizen journalism and public participation in the era of new media in Indonesia: from street to tweet," *Media and Communication*, vol. 7, no. 3, pp. 79–90, 2019.

[6] J. Wallace, "Modelling contemporary gatekeeping: the rise of individuals, algorithms and platforms in digital news dissemination," *Digital Journalism*, vol. 6, no. 3, pp. 274–293, 2018.

[7] J. Pei, K. Zhong, and J. Li, "ECNN: evaluating a cluster-neural network model for city innovation capability," *Neural Computing and Applications*, pp. 1–13, 2021.

[8] C. Orellana-Rodriguez and M. T. Keane, "Attention to news and its dissemination on twitter: a survey," *Computer Science Review*, vol. 29, pp. 74–94, 2018.

[9] Y. Tsfati, H. G. Boomgaarden, J. Strömbäck, R. Vliegenthart, A. Damstra, and E. Lindgren, "Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis," *Annals of the International Communication Association*, vol. 44, no. 2, pp. 157–173, 2020.

[10] P. Rajapaksha, R. Farahbakhsh, and N. Crespi, "Scrutinizing news media cooperation in facebook and twitter," *IEEE Access*, pp. 1–14, 2021.

[11] L. Cosgrove, I. A. Cristea, A. F. Shaughnessy, B. Mintzes, and F. Naudet, "Digital aripiprazole or digital evergreening? a systematic review of the evidence and its dissemination in the scientific literature and in the media," *BMJ Evidence-Based Medicine*, vol. 24, no. 6, pp. 231–238, 2019.

[12] S. V. Bryukhovetskaya, K. A. Artamonova, A. A. Gibadullin, S. A. Ilminskaya, and Z. M. Kurbonova, "Management of digital technology development in the national economy," *IOP Conference Series: Earth and Environmental Science*, vol. 421, no. 4, Article ID 042018, 2020.

[13] P. E. D. Love and J. Matthews, "The "how" of benefits management for digital technology: from engineering to asset management," *Automation in Construction*, vol. 107, Article ID 102930, 2019.

[14] A. Orben and A. K. Przybylski, "The association between adolescent well-being and digital technology use," *Nature Human Behaviour*, vol. 3, no. 2, pp. 173–182, 2019.

[15] A. Marusin, A. Marusin, and T. Ablyazov, "Transport infrastructure safety improvement based on digital technology implementation," in *Proceedings of the International Conference on Digital Technologies in Logistics and Infrastructure*, pp. 348–352, St. Petersburg, Russia, 2019.

[16] D. Bell, J. Wang, and S. Li, "The axial age of technology," *Issues of Contemporary World Socialism*, vol. 2, pp. 50–71, 2003.

[17] W. K. Hocking, B. Fuller, and B. Vandepeer, "Real-time determination of meteor-related parameters utilizing modern digital technology," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 63, no. 2-3, pp. 155–169, 2001.

[18] K. Healy, "Digital technology and cultural goods," *The Journal of Political Philosophy*, vol. 10, no. 4, pp. 478–500, 2002.

[19] V. Tafor, "Digital technology-understanding the problems posed by information technology in generating and managing records from a third world perspective," *ESARBICA Journal*, vol. 22, p. 72, 2003.

[20] S. Aupers, "The revenge of the machines: on modernity, digital technology and animism," *Asian Journal of Social Science*, vol. 30, no. 2, pp. 199–220, 2002.

[21] N. B. Adams, "Digital intelligence fostered by technology," *Journal of Technology Studies*, vol. 30, no. 2, pp. 93–97, 2004.

[22] M. A. Farahat, F. A. Farag, and H. A. Elsimary, "Only digital technology analog-to-digital converter circuit," in *Proceedings of the Midwest Symposium on Circuits and Systems*, vol. 1, pp. 178–181, Cairo, Egypt, 2003.

[23] M. Macedonia, "Revitalizing museums with digital technology," *Computer*, vol. 36, no. 2, pp. 94–96, 2003.

[24] W. Li-Xin, "Digital mine and the future development of mines in China," *Science & Technology Review*, vol. 7, 2004.

[25] X. Dai, "Towards a digital economy with Chinese characteristics?" *New Media & Society*, vol. 4, no. 2, pp. 141–162, 2002.

[26] Z. Jinyi, "Bicentennial challenge: speculations on China digital library construction practice," *Library and Information Service*, vol. 45, no. 3, p. 20, 2001.

[27] K. Hu, "The power of circulation: digital technologies and the online Chinese fans of Japanese TV drama," *Inter-Asia Cultural Studies*, vol. 6, no. 2, pp. 171–186, 2005.

[28] C. Jing and P. Zhong, "The study of film and TV cultural creative industry based on digital technology," in *Proceedings of the International Conference on Environmental Science and Information Application Technology*, vol. 1, pp. 705–707, Wuhan, China, 2009.

[29] M. Tang, L. Jorba, and M. J. Jensen, *Digital Media and Political Engagement Worldwide: A Comparative Study*, Cambridge University Press, Cambridge, UK, 2012.

[30] J. Liu, "Digital media, cycle of contention, and sustainability of environmental activism: the case of anti-PX protests in China," *Mass Communication and Society*, vol. 19, no. 5, pp. 604–625, 2016.

[31] S. Harlow and L. Guo, "Will the revolution be tweeted or facebooked? using digital communication tools in immigrant activism," *Journal of Computer-Mediated Communication*, vol. 19, no. 3, pp. 463–478, 2014.

[32] S. Fernandez-Lores, N. Crespo-Tejero, and R. Fernández-Hernández, "Driving traffic to the museum: the role of the digital communication tools," *Technological Forecasting and Social Change*, vol. 174, Article ID 121273, 2022.

[33] D. E. Rumelhart, "Learning representation by BP errors," *Nature*, vol. 7, pp. 64–70, 1986.

Research Article

An Integrative Bioinformatics Analysis of the Potential Mechanisms Involved in Propofol Affecting Hippocampal Neuronal Cells

Zhao Zhuang,[1] Dajiang Li,[2] Mengmeng Jiang,[2] Ye Wang,[2] Qianqian Cao,[2] Shenfeng Li,[3] Ruixue Luan,[1] Lina Sun [iD],[1] and Shoushi Wang [iD][2]

[1]*Academy of Anesthesiology, Weifang Medical University, Weifang, China*
[2]*Qingdao Central Hospital, Central Hospital Affiliated to Qingdao University, Qingdao, China*
[3]*Academy of Anesthesiology, Qingdao University, Qingdao, China*

Correspondence should be addressed to Lina Sun; sun.lina@outlook.com and Shoushi Wang; shoushiwang0502@163.com

The aim of this study is to probe the possible molecular mechanisms underlying the effects of propofol on HT22 cells. HT22 cells treated with different concentrations were sequenced, and then the results of the sequencing were analyzed for dynamic trends. Expression pattern clustering analysis was performed to demonstrate the expression of genes in the significant trend modules in each group of samples. We first chose the genes related to the trend module for WGCNA analysis, then constructed the PPI network of module genes related to propofol treatment group, and screened the key genes. Finally, GSEA analysis was performed on the key genes. Overall, 2,506 genes showed a decreasing trend with increasing propofol concentration, and 1,871 genes showed an increasing trend with increasing propofol concentration. WGCNA analysis showed that among them, turquoise panel genes were negatively correlated with propofol treatment, and genes with Cor R >0.9 in the turquoise panel were selected for PPI network construction. The MCC algorithm screened a total of five key genes (CD86, IL10RA, PTPRC, SPI1, and ITGAM). GSEA analysis showed that CD86, IL10RA, PTPRC, SPI1, and ITGAM are involved in the PRION_DISEASES pathway. Our study showed that propofol sedation can affect mRNA expression in the hippocampus, providing new ideas to identify treatment of nerve injury induced by propofol anesthesia.

1. Introduction

Propofol is one of the intravenous hypnotic medicines used to induce and maintain sedation and general anesthesia. It exerts its action through potentiation of the inhibitory neurotransmitter γ-aminobutyric acid at the GABA-A receptor [1, 2]. It is highly lipophilic and thus can rapidly cross the blood-brain barrier, leading to early onset of action. The characteristics of propofol are well known. It has fast metabolism and short recovery time, regardless of the depth or length of sedation time [3]. In addition to being used as an anesthetic, researchers also found that propofol is also related to many cancer-related pathophysiological processes and can play an important role in cancer by regulating the expression of a variety of downstream molecules, long-chain noncoding RNA, microRNA, and signal pathways [4]. Propofol upregulates miR-195, and then JAK/stat and NF-κB pathway is inactivated to inhibit the proliferation, migration, and invasion of gastric carcinoma MKN45 cells [5].

Although propofol can play an excellent role in maintaining sedation and general anesthesia, studies have also focused on changes in regional cerebral blood flow or neuronal activity during propofol sedation. Only one study showed that as the depth of propofol sedation increased, the activity in the area corresponding to the stimulus applied during propofol sedation decreased. In addition, they found that propofol sedation impaired the function of basal ganglia circuits and thalamocortical connections [6]. It is
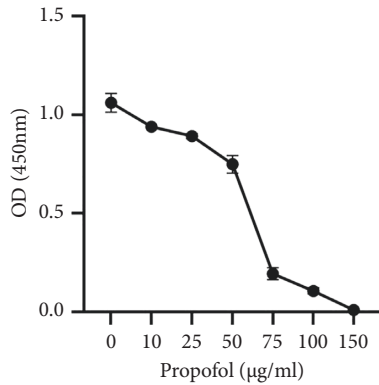
FIGURE 1: CCK8 shows the processing effect.

understood that learning and memory occur mainly in the hippocampus, and propofol may cause hippocampal neurotoxicity [7, 8]. General anesthetics may induce developmental neurotoxicity, including subsequent long-term memory, acute extensive nerve cell death, and abnormal learning behavior [9]. Although many researchers have figured out that propofol induced neurotoxicity in the brains of developing animals, its exact mechanism of action remains largely unknown. Then, this manuscript, on the other hand, was conducted to observe the potential biological processes and signaling pathways that may play an important role during propofol sedation through the biology information analysis of genes with different expression patterns.

2. Methods

2.1. Cell Processing and Sequencing.
The HT22 hippocampal cell line used in the present study was purchased from Shanghai Yaji Biotechnology Co. (Shanghai, China). The cells were fostered in DMEM containing 10% FBC, $100 \mu/\text{mL}$ penicillin, 100ng/mL streptomycin at 37°C, and 5% CO_2. When the cell density reached 60–70%, the HT22 cells were exposed to propofol at concentrations of $25 \mu\text{g/ml}$, $50 \mu\text{g/ml}$, $75 \mu\text{g/ml}$, and $100 \mu\text{g/ml}$ for 1 day. The experiment was repeated three times. The cDNA/DNA/Small RNA was sequenced on the Illumina sequencing platform by Genednovo Biotechnology (Guangzhou, China). The libraries were sequenced.

2.2. Effects of Treatment on CCK-8.
Cells were inoculated in 96-well plates at $1 \times 10\ 3$ cells per well. Cell Counting Kit-8 (CCK-8; Beyotime Biotechnology, Shanghai, China) assays were performed daily. Briefly, the cells were incubated in a $100 \mu\text{L}$ medium containing $10 \mu\text{L}$ of CCK-8 (0.5mg/mL) reagent at 37°C for one to four hours. Absorbance was measured at 450 nm by an enzyme-labeled assay (Figure 1). The experiment was performed at least three times in three replicate wells.

2.3. Trend Analysis.
The results obtained from sequencing were subjected to follow-up experiments. The expression of each gene was grouped according to different concentration

points, and then a dynamic trend analysis was completed using OmicShare Tools. A P value <0.05 would show a significant difference.

2.4. Protein-Protein Interaction (PPI) Network Construction.
All associations got from STRING were provided with a confidence score and imported into Cytoscape software. Modules of the PPI network were screened using the Molecular Complex Detection plugin in Cytoscape. A P value <0.05 would show a significant difference.

2.5. Gene Ontology (GO) Function.
Gene ontology annotations and KEGG pathway enrichment analyses were performed using the Enrichr database GO terms consisting of the following three components: biological process (BP), cellular component (CC), and molecular function (MF).

2.6. WGCNA Analysis.
Co-expression networks were constructed using the WGCNA data bank in R software. Paired Pearson correlations were first used to evaluate the weighted co-expression relationships between all subjects in the dataset in the adjacency matrix. The matrices were then converted to TOMs using the topological overlap matrix (TOM) similarity function, and the resulting TOMs were used to measure co-expression relationships between genes based on biologically meaningful genetic similarity.

2.7. GSEA Analysis.
The genes were classified into high and low-expression parts, and then GSEA V3.0 software was used to analyze the enrichment results for the genes. A nominal P value of <0.05 and false discovery rate (FDR) of <25% were selected as cut-off criteria. We selected the top two ranked analysis results.

3. Results

3.1. Trend Analysis.
We first subjected the expression data obtained from sequencing to trend analysis to observe the genes whose expression continued to increase or decrease with increasing propofol concentration. The results of the trend analysis showed that genes concentrated in module 0 decreased in expression with increasing propofol concentration, and genes concentrated in module 19 increased in expression with increasing propofol concentration; both panels were differentially significant. There were 2,506 genes in module 0 and 1,871 genes in module 19 (see Figure 2).

3.2. Heat Map Showing the Expression of Each Gene in Each Sample.
Subsequently, we selected statistically significant modules for expression pattern clustering analysis. The heat map shows the expression of each trend module gene in each group of samples (see Figure 3).

3.3. WGCNA Analysis.
We selected module 0, in which gene expression was downregulated with propofol concentration, and module 19, in which gene expression increased with
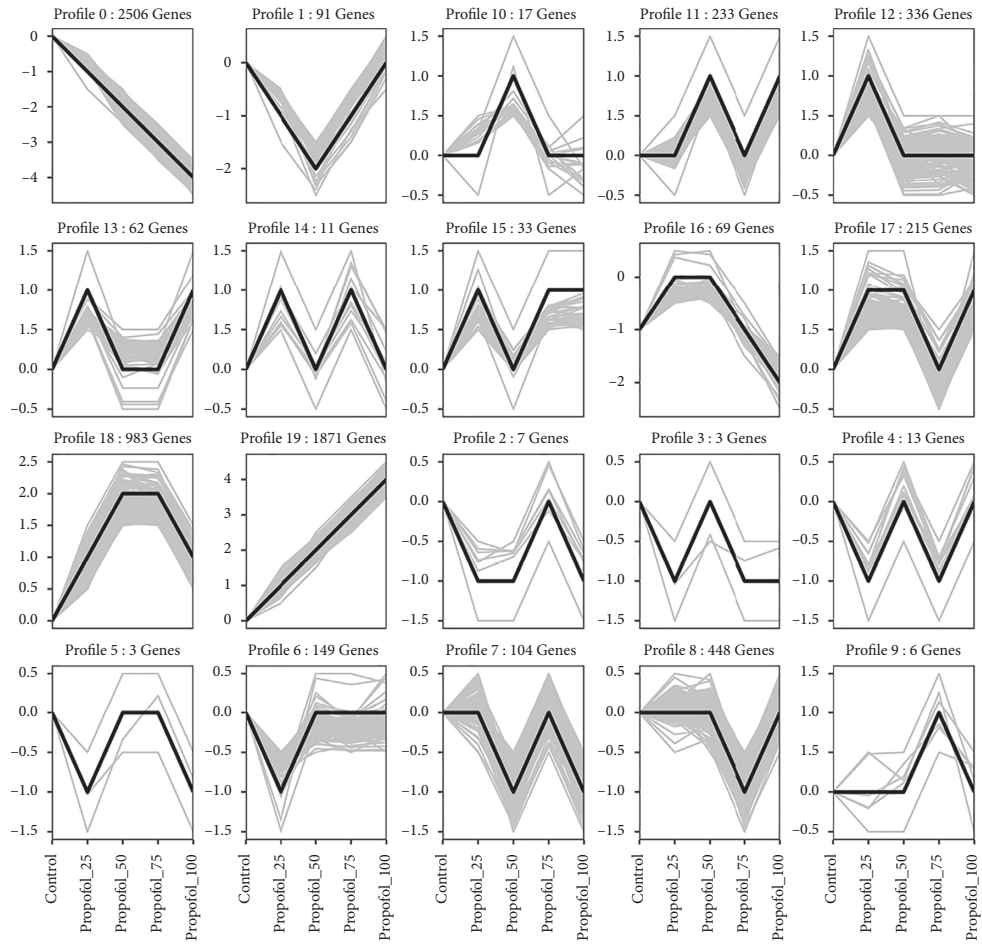
(a)

Profile ordered based on the pvalue significance of number of Genes assigned versus expected

(b)

Figure 2: Continued.

(c)
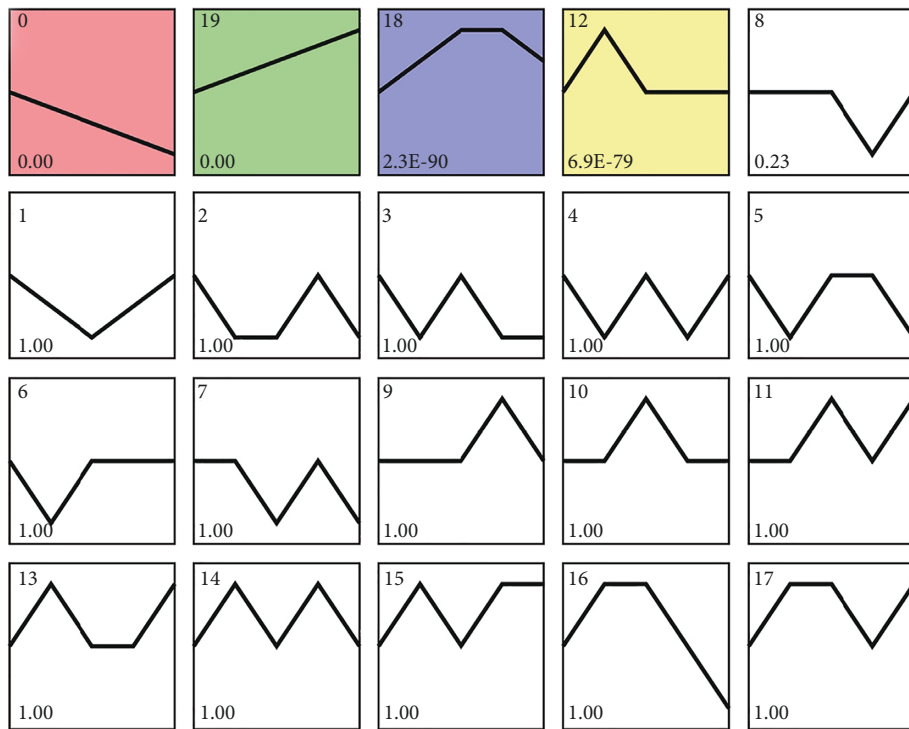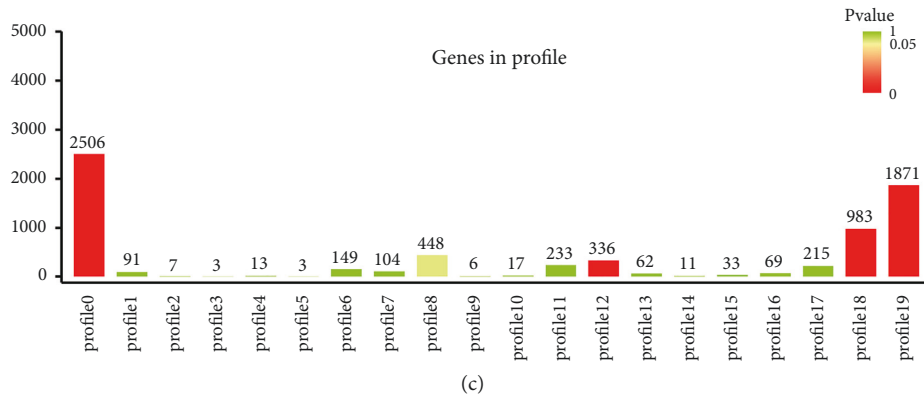
FIGURE 2: Trend analysis. (a) Number of genes in each module. (b) P-value of each module. (c) Bar graph showing the number of genes and P-value of each module.

propofol concentration; we collected genes from both modules for WGCNA analysis. The analysis identified a total of two panels, in which the turquoise panel genes were negatively correlated with the treatment of propofol (see Figure 4).

3.4. PPI Network Construction.
We selected the genes with Cor R >0.9 in the turquoise panel, a total of 1,698 genes, to construct the PPI network. Subsequently, the top 5 key genes, i.e., CD86, IL10RA, PTPRC, SPI1, and ITGAM, were screened using the MCC algorithm (see Figure 5).

3.5. GSEA Analysis.
We grouped the samples as high and low parts, while GSEA analysis was performed on single genes afterward. The analysis figured out that the low expression of CD86, IL10RA, PTPRC, SPI1, and ITGAM all have important roles in the PRION_DISEASES pathway, and CD86, PTPRC, SPI1, and ITGAM are also all involved in the NUCLEOTIDE_EXCISION_REPAIR signaling pathway. Furthermore, IL10RA is involved in the CIRCA-DIAN_RHYTHM_MAMMAL signaling pathway (see Figure 6).

4. Discussion

Propofol causes fewer anesthetic side effects and facilitates faster recovery compared to other intravenous anesthetics. The benefits of propofol include the rapid onset of anesthesia, short recovery time, and neuroprotective effects in pathogenic situations [8, 10]. Hence, concerns have been raised about the possible effects of the widespread use of propofol on the central nervous system. Propofol does have neurotoxic effects on children. It affects brain development by inhibiting neuronal activation of hippocampal neurons, which may lead to the reduction of neurocognitive function [11]. Animal experiments have shown that exposure to subanesthetic doses of isoproterenol alters the long non-coding RNA profile in the hippocampus of immature mice and leads to disruption of hippocampal circuits, while

exposure to high doses of isoproterenol inhibits long-term potentiation in the CA1 region of the adult hippocampus [12]. Previous studies have suggested that this may be because propofol is the most potent drug for activating the GABA-A currents in the immature hypothalamus [13]. To date, the specific molecular mechanisms involved in propofol affecting the hippocampal neuronal cells still remain unclear. To find the potential mechanisms by which propofol may affect hippocampal neuronal cells, we performed a trend analysis of the results obtained from sequencing. We found that 2,506 genes may decrease in expression with increasing propofol concentrations, and 1,871 genes may increase in expression with increasing propofol concentrations. We then subjected these 4,377 genes to WGCNA analysis and found that the turquoise plate genes were significantly and negatively correlated with the propofol in your treatment group. We then selected the genes in this panel with correlation coefficients greater than 0.9 for PPI network construction and screened five key genes: CD86, IL10RA, PTPRC, SPI1, and ITGAM.

CD86 is a type I transmembrane protein associated with T-cell activation in the immune system [14]. CD86 may act as a key regulator of the immune response to disease through a T cell-mediated mechanism, and thus it has great potential to become a new target for immunotherapy [15]. IL10RA consists mainly of a heterotetramer of the anti-inflammatory cytokine IL10 receptor and belongs to class II cytokines [16, 17]. IL10RA is usually located as a cell surface receptor upstream of STAT3, and it can bind IL-10 together with IL10RB to mediate downstream signaling via STAT3 [18]. PTPRC is expressed on all nucleated cells of the hematopoietic system. It expresses several isoforms specific to a certain cell type and cell development or activation state [19]. PTPRC also plays an important role in autoimmune diseases and cancers, as well as in infectious diseases; its deficiency may lead to T and B lymphocyte dysfunction, manifesting as severe combined immunodeficiency [20].

Spi1 is a major regulator of hematopoiesis because it is involved in the self-renewal of hematopoietic stem and progenitor cells, as well as in the stereotyping and
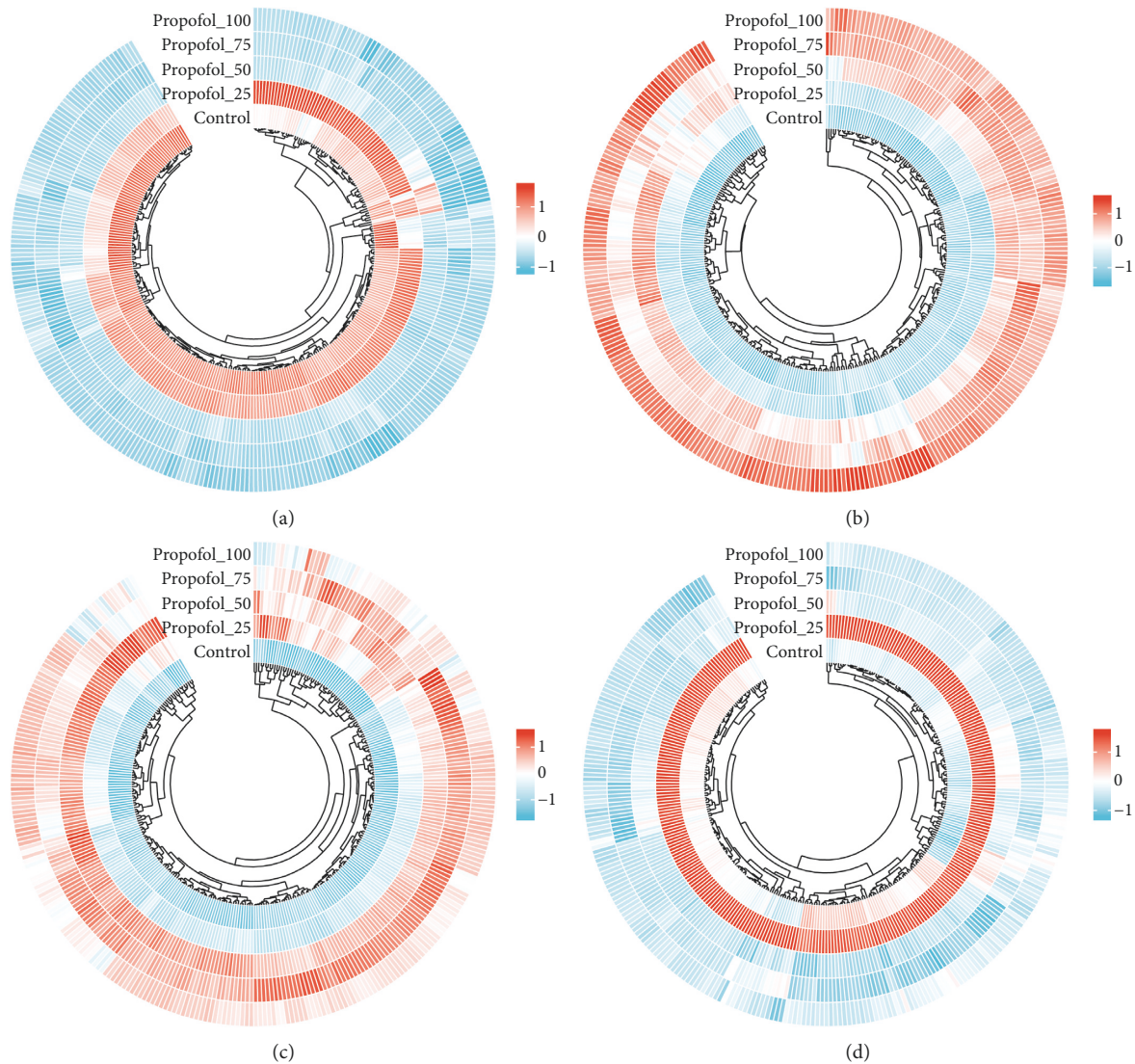
FIGURE 3: Heat map showing the expression of each gene in each sample. (a) Heatmap of module 0 gene expression in each sample. (b) Heatmap of module 19 gene expression in each sample. (c) Heatmap of module 18 gene expression in each sample. (d) Heatmap of module 12 gene expression in each sample.

maturation of the myeloid and B-lymphatic lineages. In-appropriate Spi1 expression is oncogenic. However, the molecular mechanisms mediating the oncogenic function of Spi1 are complex and not yet fully understood [21]. In addition, Spi1 transcription factors are important in regulating macrophage and neutrophil development [22]. ITGAM is expressed by a variety of myeloid cell types, which can form heterodimers with CD18 and mediate adhesion between cell types in the immune system [23, 24]. ITGAM deficiency may enhance disease progression and inflammation in multiple autoimmune models, including lupus [25]. The above then could show that CD86, IL10RA, PTPRC, SPI1, and ITGAM all play important roles in immune response. Inflammation is a response triggered by damage to living tissue. It can be beneficial or may lead to tissue destruction. Inflammation is particularly dangerous when the nervous system is involved (so-called

"neuroinflammation"), whether acute or chronic [26]. In turn, neuroinflammation can cause reversible and irre-versible neurological sequelae, including cognitive impair-ment [27]. However, by modulating inflammation, it is also possible to improve peripheral nerve repair [28]. It has been suggested that the abnormal expression of CD86, IL10RA, PTPRC, SPI1, and ITGAM may trigger an immune-in-flammatory response and thus lead to neurological damage.

In the present experiment, we performed GSEA analyses of individual genes separately and showed that CD86, IL10RA, PTPRC, SPI1, and ITGAM are involved in the PRION_DISEASES pathway. Prion diseases can make the prion protein (PrP Sc) in the central nervous system itch. In the early stages, microglia respond to the deposition of PrP Sc, thereby increasing their phagocytic capacity to clear PrP Sc. However, this phagocytosis is not sufficient and can have detrimental effects on the brain [29]. This is further evidence

(a)

(b)
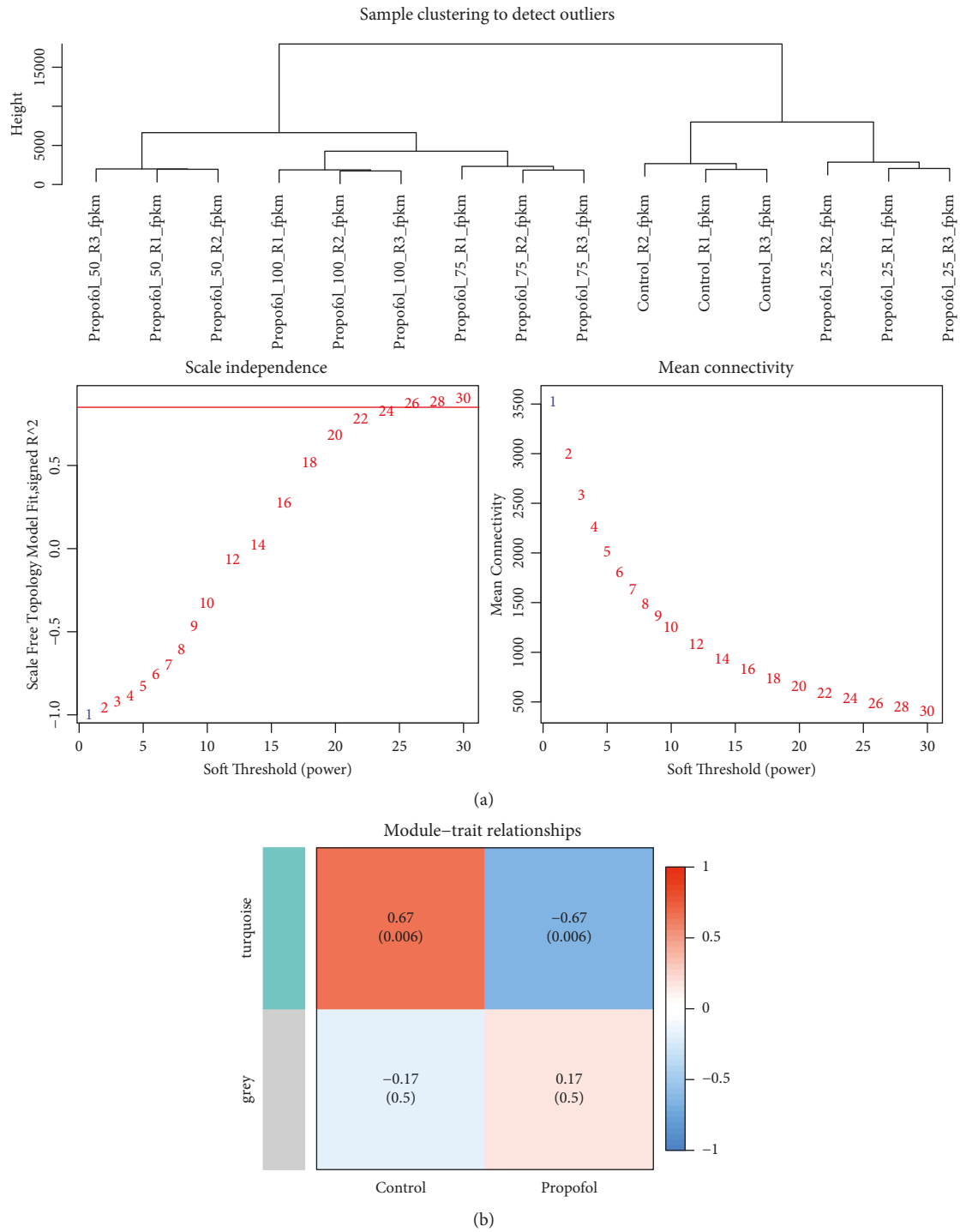
FIGURE 4: WGCNA analysis. (a) Heat map of correlations between (a) soft threshold (b) module signature genes and propofol treatment groups.

(a)

(b)
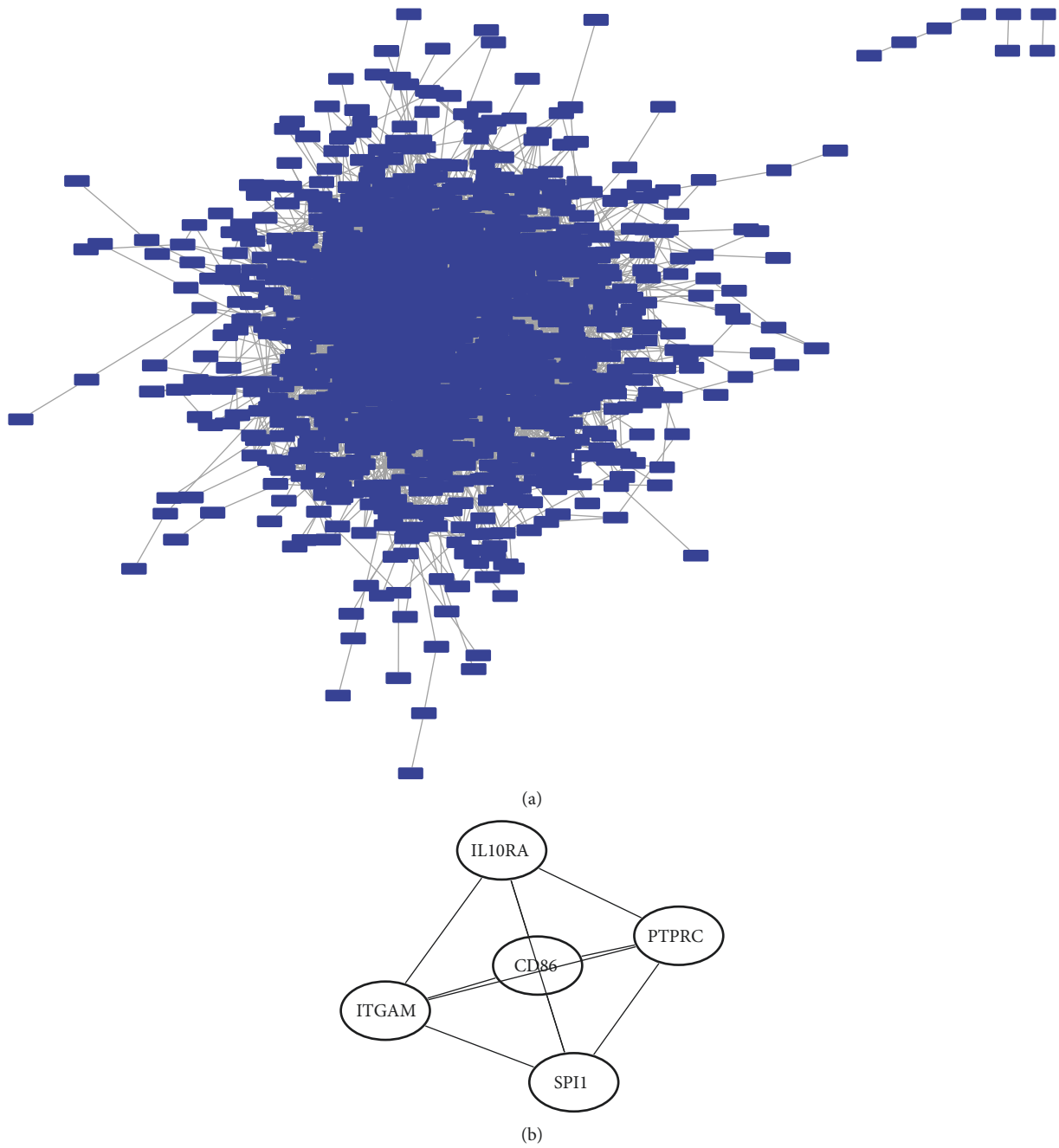
FIGURE 5: PPI network construction. (a) PPI network building block diagram. (b) 5 key genes obtained from MCC algorithm analysis.

(a)

(b)

(c)
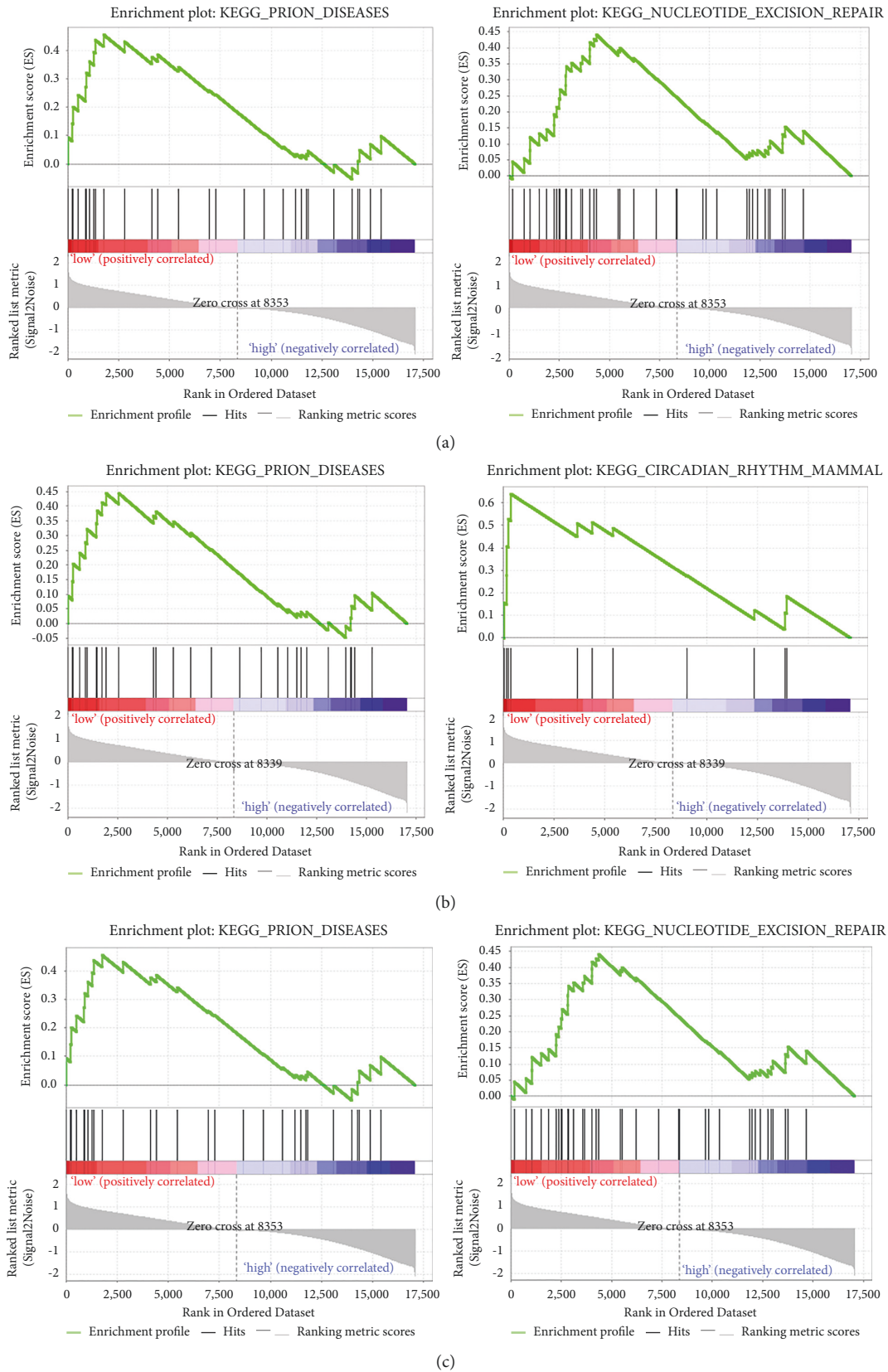
FIGURE 6: Continued.

(d)

(e)
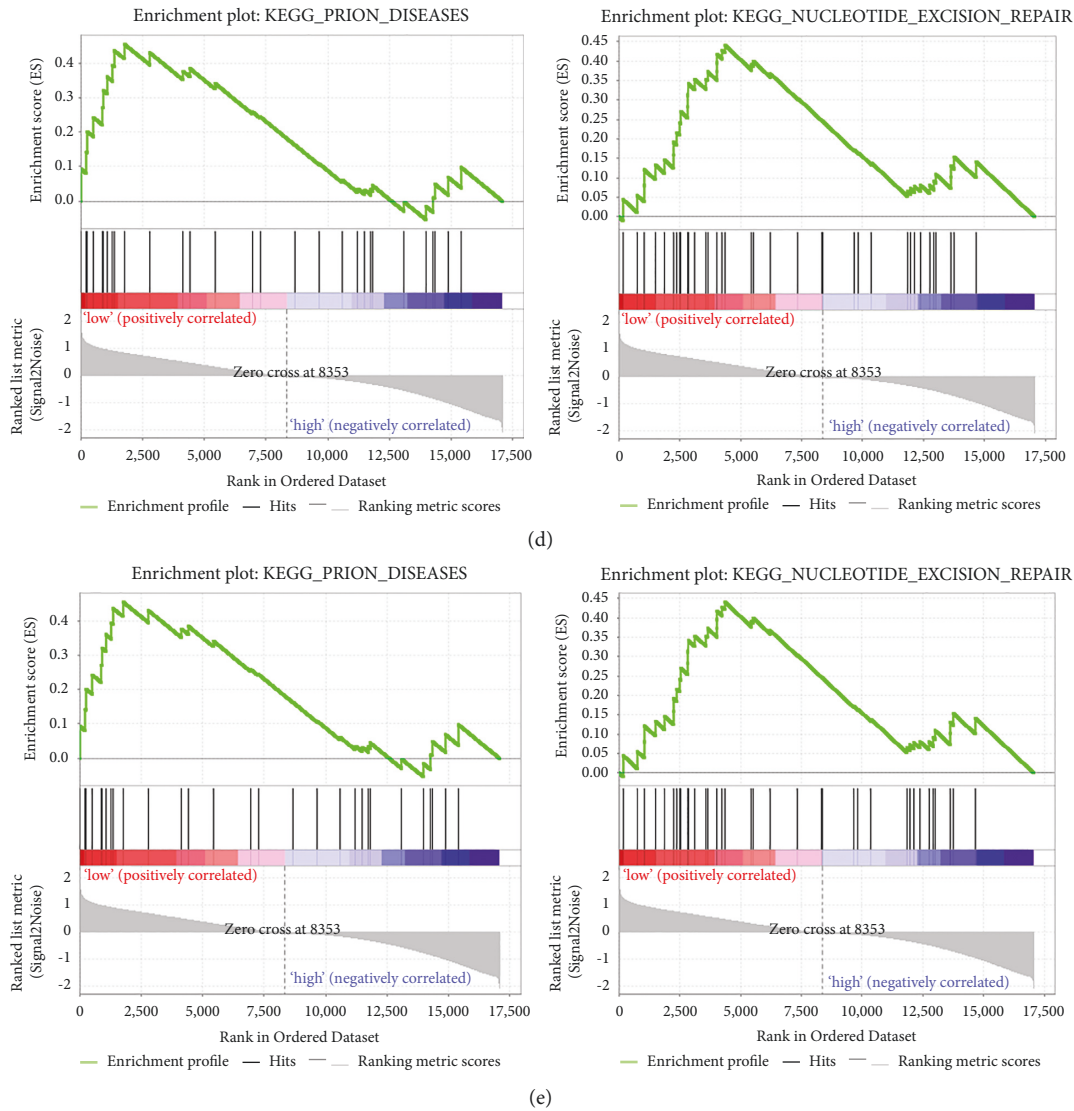
FIGURE 6: GSEA analysis. (a) The first two signaling pathways involved in CD86 low expression. (b) The first two signaling pathways involved in IL10RA low expression. (c) The first two signaling pathways involved in PTPRC low expression. (d) The first two signaling pathways involved in SPI1 low expression. (e) The first two signaling pathways involved in ITGAM low expression.

that CD86, IL10RA, PTPRC, SPI1, and ITGAM have an important relationship in neurological impairment of the brain. This could also suggest that propofol may lead to neurological impairment by affecting the abnormal expression of CD86, IL10RA, PTPRC, SPI1, and ITGAM; thus, the exact mechanism remains to be investigated.

5. Conclusion

In conclusion, CD86, IL10RA, PTPRC, SPI1, and ITGAM may play an important role in propofol affecting mouse hippocampal neuronal cells. Propofol may affect mouse hippocampal neurons by affecting the expression of PTPRC, CD86, il10ra, spi1il10ra and ITGAM. Thereby, these findings may imply potential new targets for the treatment of

patients with propofol anesthesia-induced neurological injury.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] M. M. Sahinovic, M. M. R. F. Struys, and A. R. Absalom, "Clinical pharmacokinetics and pharmacodynamics of propofol," *Clinical Pharmacokinetics*, vol. 57, no. 12, pp. 1539–1558, 2018.

[2] L. Chen, Z.-l. Yang, J. Cheng et al., "Propofol decreases the excitability of cholinergic neurons in mouse basal forebrain via GABAA receptors," *Acta Pharmacologica Sinica*, vol. 40, no. 6, pp. 755–761, 2019.

[3] T. Nishizawa and H. Suzuki, "Propofol for gastrointestinal endoscopy," *United European Gastroenterology Journal*, vol. 6, no. 6, pp. 801–805, 2018.

[4] Y. Xu, S. Pan, W. Jiang, F. Xue, and X. Zhu, "Effects of propofol on the development of cancer in humans," *Cell Proliferation*, vol. 53, no. 8, Article ID e12867, 2020.

[5] W. Zhang, Y. Wang, Z. Zhu, Y. Zheng, and B. Song, "Retracted: propofol inhibits proliferation, migration and invasion of gastric cancer cells by up-regulating microRNA-195," *International Journal of Biological Macromolecules*, vol. 120, no. A, pp. 975–984, 2018.

[6] R. N. Mhuircheartaigh, D. Rosenorn-Lanng, R. Wise, S. Jbabdi, R. Rogers, and I. Tracey, "Cortical and subcortical connectivity changes during decreasing levels of consciousness in humans: a functional magnetic resonance imaging study using propofol," *Journal of Neuroscience*, vol. 30, no. 27, pp. 9095–9102, 2010.

[7] S. Zhang, Z. Liang, W. Sun, and L. Pei, "Repeated propofol anesthesia induced downregulation of hippocampal miR-132 and learning and memory impairment of rats," *Brain Research*, vol. 1670, pp. 156–164, 2017.

[8] X. Li, L. Yao, Q. Liang, H. Qu, and H. Cai, "Propofol protects hippocampal neurons from hypoxia-reoxygenation injury by decreasing calcineurin-induced calcium overload and activating YAP signaling," *Oxidative Medicine and Cellular Longevity*, vol. 2018, Article ID 1725191, 12 pages, 2018.

[9] L. Feng, Z. G. Sun, Q. W. Liu et al., "Propofol inhibits the expression of Abelson nonreceptor tyrosine kinase without affecting learning or memory function in neonatal rats," *Brain and behavior*, vol. 10, no. 11, Article ID e01810, 2020.

[10] N. Berndt, O. Haq, O. Kovács et al., "Possible neurotoxicity of the anesthetic propofol: Evidence for the inhibition of complex II of the respiratory chain in area CA3 of rat hippocampal slices," *Archives of Toxicology*, vol. 92, no. 10, pp. 3191–3205, 2018.

[11] Y. Yang, J. Yi, M. Pan, B. Hu, and H. Duan, "Edaravone alleviated propofol-induced neurotoxicity in developing Hippocampus by mBDNF/TrkB/PI3K pathway," *Drug Design, Development and Therapy*, vol. 15, pp. 1409–1422, 2021.

[12] R. Guan, J. Lv, F. Xiao, Y. Tu, Y. Xie, and L. Li, "Potential role of the cAMP/PKA/CREB signalling pathway in hypoxic preconditioning and effect on propofolinduced neurotoxicity in the hippocampus of neonatal rats," *Molecular Medicine Reports*, vol. 20, no. 2, pp. 1837–1845, 2019.

[13] V. Tesic, S. M. Joksimovic, N. Quillinan et al., "Neuroactive steroids alphaxalone and CDNC24 are effective hypnotics and potentiators of GABAA currents, but are not neurotoxic to the developing rat brain," *British Journal of Anaesthesia*, vol. 124, no. 5, pp. 603–613, 2020.

[14] J. Fang, H. Ying, T. Mao et al., "Upregulation of CD11b and CD86 through LSD1 inhibition promotes myeloid differentiation and suppresses cell proliferation in human monocytic leukemia cells," *Oncotarget*, vol. 8, no. 49, pp. 85085–85101, 2017.

[15] H. Qiu, W. Tian, Y. He et al., "Integrated analysis reveals prognostic value and immune correlates of CD86 expression in lower grade glioma," *Frontiers in Oncology*, vol. 11, Article ID 654350, 2021.

[16] F. A. Al-Abbasi, K. Mohammed, S. Sadath, B. Banaganapalli, K. Nasser, and N. A. Shaik, "Computational protein phenotype characterization of IL10RA mutations causative to early onset inflammatory bowel disease (IBD)," *Frontiers in Genetics*, vol. 9, p. 146, 2018.

[17] A.-J. Xue, S.-J. Miao, H. Sun et al., "Intestinal dysbiosis in pediatric Crohn's disease patients with IL10RA mutations," *World Journal of Gastroenterology*, vol. 26, no. 22, pp. 3098–3109, 2020.

[18] N. Prokoph, N. A. Probst, L. C. Lee et al., "IL10RA modulates crizotinib sensitivity in NPM1-ALK+ anaplastic large cell lymphoma," *Blood*, vol. 136, no. 14, pp. 1657–1669, 2020.

[19] J. Landskron, S. M. Kraggerud, E. Wik et al., "C77G in PTPRC (CD45) is no risk allele for ovarian cancer, but associated with less aggressive disease," *Plos One*, vol. 12, no. 7, Article ID e0182030, 2017.

[20] A. Rheinländer, B. Schraven, and U. Bommhardt, "CD45 in human physiology and clinical medicine," *Immunology Letters*, vol. 196, pp. 22–32, 2018.

[21] P. Rimmelé, M. Esposito, L. Delestré et al., "The Spi1/PU.1 transcription factor accelerates replication fork progression by increasing PP1 phosphatase in leukemia," *Oncotarget*, vol. 8, no. 23, pp. 37104–37114, 2017.

[22] A. Zakrzewska, C. Cui, O. W. Stockhammer, E. L. Benard, H. P. Spaink, and A. H. Meijer, "Macrophage-specific gene functions in Spi1-directed innate immunity," *Blood*, vol. 116, no. 3, pp. e1–e11, 2010.

[23] G. Hom, R. R. Graham, B. Modrek et al., "Association of systemic lupus erythematosus withC8orf13-BLKandITGAM-ITGAX," *New England Journal of Medicine*, vol. 358, no. 9, pp. 900–909, 2008.

[24] J. L. Dunne, R. G. Collins, A. L. Beaudet, C. M. Ballantyne, and K. Ley, "Mac-1, but not LFA-1, uses intercellular adhesion molecule-1 to mediate slow leukocyte rolling in TNF-α-induced inflammation," *The Journal of Immunology*, vol. 171, no. 11, pp. 6105–6111, 2003.

[25] J. Ramirez-Bello, C. Sun, G. Valencia-Pacheco et al., "ITGAM is a risk factor to systemic lupus erythematosus and possibly a protection factor to rheumatoid arthritis in patients from Mexico," *Plos One*, vol. 14, no. 11, Article ID e224543, 2019.

[26] S. D. Skaper, L. Facci, M. Zusso, and P. Giusti, "An inflammation-centric view of neurological disease: beyond the neuron," *Frontiers in Cellular Neuroscience*, vol. 12, p. 72, 2018.

[27] A. Alam, Z. Hana, Z. Jin, K. C. Suen, and D. Ma, "Surgery, neuroinflammation and cognitive impairment," *EBioMedicine*, vol. 37, pp. 547–556, 2018.

[28] T.-C. Huang, H.-L. Wu, S.-H. Chen, Y.-T. Wang, and C.-C. Wu, "Thrombomodulin facilitates peripheral nerve regeneration through regulating M1/M2 switching," *Journal of Neuroinflammation*, vol. 17, no. 1, p. 240, 2020.

[29] A. Aguzzi and C. Zhu, "Microglia in prion diseases," *Journal of Clinical Investigation*, vol. 127, no. 9, pp. 3230–3239, 2017.

Research Article

Research on the Method of Digital Media Content Creation Based on the Internet of Things

Yiming Sun [ID]

Fine Arts Academy, Changchun Normal University, Changchun, Jilin 130031, China

Correspondence should be addressed to Yiming Sun; sunyiming@ccsfu.edu.cn

This paper provides an in-depth study and analysis of the approach to digital media content creation using IoT technology. This study first investigates both WSN and NB-IoT technologies separately and then combines them to investigate the networking structure based on NB-IoT and WSN as well as the related technologies based on fusion. Then, the coverage method of the node redeployment of the traditional network of wireless sensors is elaborated. For the traditional method, a series of problems such as poor node connectivity and low coverage caused by not conducting a local search of subgroup nodes are proposed, and the node redeployment method of wireless sensor network based on the frog hopping algorithm is proposed, introducing the frog hopping algorithm, combining local search of subgroup and global information exchange, and proposing that for wireless sensor network nodes. We propose initializing the nodes of the wireless sensor network, combining the global information exchange with the subgroup, determining the range of the detectable area of the wireless sensor, and solving the calculus method for the irregular area to get the optimal network nodes so that the nodes of the wireless sensor network can be redeployed. In terms of system, this design abandons the CIS structure, which is more used in the current IoT and uses B/S structure and Java-web technology to design a web-based system. SSH (Struts + Spring + Hibernate) framework, AJAX asynchronous request technology, and JXLL/ JFreechart technology are fully applied to improve the practicality and comfort of the system, reduce development and maintenance costs, and facilitate popularization. After the system design was completed, the modules were tested with the existing experimental conditions. The experimental results show that the system is designed to be reliable and can operate normally with high stability and practical significance.

1. Introduction

The Internet of Things (IoT) is an extension and expansion of the Internet and communication networks, which mainly identifies the physical world through intelligent devices and sensing technologies, and carries out knowledge mining, processing, and computing based on network transmission and interconnection, to achieve the goal of seamless linkage and information exchange between things and things, and things and people, and thus accurately manage and control the physical world in real time based on scientific decision-making [1]. Since then, the Internet of Things (IoT) has received increasing attention from countries around the world, and the current consensus is that IoT belongs to another round of information technology achievements after mobile communication networks, the Internet, and

computer, which is extremely disruptive and revolutionary [2]. Former US President and the first black president Barack Obama responded to the "Smart Planet" strategy (proposed by IBM) as early as 2009, and during his term of office, the IoT strategy was fully integrated into the US national development strategy. The EU IoT Action Plan is the world's first IoT development strategy, which was officially released by the EU in 2009 and describes the trends in IoT technology development and application as its basic content [3]. Advances in science and technology have created new platforms and spaces for artists, greatly enriching the forms of artistic creation and promoting the development of art [4]. The rapid development of technology and media has led to a surprising collision between digital media technology and art in the era of mobile intelligence, and art creation has since had more new forms of expression and development

platforms, while the reading needs of the general audience have given art unlimited creative impetus and inspiration [4]. It is widely used in various exhibition halls, digital display, 3D virtual display, interactive display, guide system design, and other aspects of the use of digital media art. In recent years, digital media art has progressed particularly fast in display design activities, and the use of digital TV, digital broadcasting, digital advertising, virtual reality technology, laser light, heat-sensing equipment, etc. has become more widespread, which plays a vital role in improving the visual effect of the display, creating atmosphere, and promoting publicity [5].

The significance of this paper is to use the performance and planning of IoT-NB coverage as a basis to lay a solid foundation for future studies on coverage in other frequency bands, and to serve as a reserve for both coverage enhancement and interference prevention in the future. IoT-NB can be considered as a simplified LTE network, which is characterized by wide coverage area and large connections, and it plays a key role in the analysis and evaluation of IoT-NB IoT coverage and the planning of network capacity and scale [6]. The simulation and calibration of the model in this paper can be used as a guide for the subsequent planning, which is very important in the process of building the network. Today, global digitization is becoming more evident, and under this premise, the need for mobile communications is relatively high, and the efficiency and capability of communications are being enhanced [7]. 5G is a technology characterized by high reliability, low latency, and low power loss, which is no longer possible with simple wireless access technologies, which means that most of the new wireless access technologies need to be integrated to provide a combined solution. Combining new wireless access technologies can provide a joint solution.

2. Related Works

Contemporary digital media technology is formed based on digital information technology. Since the process of informatization of Western society started earlier, the digitization of media forms was earlier, Western scholars began to explore digital media technology from various fields and angles, and formed relatively comprehensive and in-depth theoretical research results [8]. From the macro perspective of digital media technology, many scholars have explored the impact of new digital media on the times and society. Marshall McLuhan proposed the concepts of "media as the message" and "media as human extension" in "Understanding Media—On Human Extension," revealing the significance of media technology in changing human culture and life, and in the age of electronic media, the limits of time and space will be broken through, and the distance between people in time and space will shrink, thus forming a "global village" [9]. With its precise analysis and prediction, this work has become an enlightening theoretical foundation in the field of digital media. In "Virtual Art: From Illusion to Immersion," I distinguish immersion in the digital media age from earlier forms of illusion art, draw on the actual works of contemporary artists and groups in the analysis, and summarize how to use 3D, IMAX, virtual reality, and other technologies. The

means create immersive illusions and outline the impact of virtual reality on contemporary art perceptions [10]. In "Digital Survival," it is described that in the world of the information age "information" is rapidly replacing material and becoming the basic exchange of society, and it explains that digital media technology based on the Internet and digital technology has an impact on human learning and work. In addition to the influence of entertainment and other lifestyle, it also has an important impact on the discussion and research on digital media technology [11].

The concept of the Internet of Things (IoT) was first proposed by Bill Gates, the president of Microsoft, but it was not given much attention at the strategic level because the development of the Internet was still in the rapid development stage and the related technology of IoT was not fully mature. Samie F took the transmission of radius sequence as the fundamental basis to express the distribution and heuristic algorithm of ant colony optimization, and then through the actual location of nodes [12]. The deployment is calculated in real-time to get the best node communication radius. This algorithm is heuristic, and although it can converge quickly to a more reasonable result, there is no way to ensure a continuous extension of the network lifetime because the nodes are the closest to the most correct solution and the algorithm is relatively difficult and costly. Many terminals, a small amount of data transmission, and low latency requirements are the characteristics of the current model of the Internet of Things. In terms of theoretical research, experts, and scholars such as Qiu et al. discussed the definition of IoT, technical directions, and the current development status of the IoT industry in China and worldwide and make a key analysis in each link of IoT industry chain and applications. The development status of the IoT application industry is also analyzed [13]. The Department of Technology and Economic Research of the Development Research Center of the State Council (2011) published a study on the development mechanism of China's IoT industry. The book presents the current situation of China's application of IoT technology and its problems in the sensor, cloud computing, smart grid, intelligent transportation, and electronic medical record industries [14]. The EPC field sees it as a mega-network covering everything in the world through next-generation information technology on the Internet. Any item in the network can interact with each other with data. In essence, it is the goal of information sharing by making objects automatically identify and transmit data through the Internet as a platform.

In summary, in terms of IoT technology, the promotion of the information age background has led to detailed and in-depth research both from a macro perspective and from a micro perspective of technology in the field of display, and on this solid foundation, there is still much room for improvement in research and practice regarding the use of technology in digital media content creation.

3. IoT-Based Digital Media Content Creation

3.1. IoT Model Design. The popularity and application of IoT technology have broken the traditional way of relying on a single information transfer between people. At the same

time, the information transmission of IoT technology still needs to rely on information sensing equipment to achieve, with the help of equipment and processing [15]. The Internet of Things is essentially an extension and expansion of the Internet and communication networks. It mainly recognizes the physical world through intelligent devices and sensing technology and carries out knowledge mining, processing, and calculation based on network transmission and interconnection.

The roles and functions of each structure of IoT technology are also different. The perception layer is mainly used for information perception, such as information collection and identification; the network layer is mainly used for transmission of various information, while the application layer mainly presents the results of the final identification information in front of human beings to further promote the intelligent management of something by human beings, as shown in Figure 1.

As the basic unit of the neural network, the single neuron has self-learning and self-adaptive capability and has a simple structure and fast response time. The combination of single neuron and traditional PID controller to form a single neuron adaptive PID controller, to a certain extent, solves the traditional PID controller and is not easy to adjust parameters online in real-time; it is difficult to effectively control complex processes and time-varying system parameters and other problems. Neural networks are characterized by learning, and the basic means of learning is learning rules. The connection weights are changed utilizing learning so that the adaptive function of the single neuron controller is realized. The learning algorithm is the core of the whole controller and w_i is the rule that adjusts the connection weights and reflects its ability to learn. The goal of seamless linking and information exchange between them is to accurately manage and control the physical world in real-time based on scientific decision-making. The learning algorithm is as follows:

$$w_i(k) = w_i(\kappa + 1)^2 - \lambda \lin(\kappa). \tag{1}$$

NB-IoT is an important branch of IoT technology, which has obvious differences from the traditional model in terms of voice service model and user distribution model. In traditional 2G, 3G, and 4G, the number of user accesses in base stations is controlled around 1000 and 400 in densely populated areas to ensure latency. However, NB-IoT does not require high latency, and it is dormant most of the time when there is no data interaction and activates the connection again when there is data exchange, which greatly improves the number of connections [16]. In terms of core network architecture, NB-IoT optimizes the EPC core network architecture, which can be known according to 3GPP's R13 protocol, and the principle of core network architecture is shown in Figure 2. Although there are certain differences in the way and focus of information transmission between the Internet and the Internet of Things, the ultimate purpose to be achieved is the same, which is to promote information transmission and communication between things and people. For the Internet of Things, this project the

birth of technology has laid a solid foundation for the further improvement and development of Internet technology. The user-level management of network airports is performed in FDD mode, and the process is implemented by first accessing the NB RAN with airports and then accessing the C-SGN (Cellular IoT Service Gateway) network at the S1 interface. The core capabilities are shared to third-party platforms using standard API interfaces. The core capabilities mainly include contract data management and location update management (HSS), terminal SMS function interoperability (SMS-GMSC/IWMSC/SMS ROUTER), and service capability development platform (SCEF), etc.

Using the HOP-DV algorithm to introduce multiple mobile anchor points, since each WSN contagion has its self-assembled graph, they can use neighboring nodes to transmit data to the aggregation node, so the most economical way is to distribute a mobile access point per cluster; this access point acts as the aggregation node and mobile anchor point. The mobile anchor point can obtain its location arbitrarily, and after receiving the information from all nodes in the cluster, it feeds back the number of hops and the shortest path, moves to other locations in turn, and repeats the above communication process. The coverage of all nodes in the cluster and the nodes are limited by the communication distance, and the distance moved by the mobile anchor point cannot be too small at each time of generating the movement but must be evenly distributed within the cluster. NB-IoT does not require high latency. When there is no data interaction, it is in a dormant state most of the time, and the connection is activated again when there is data exchange. This design greatly increases the number of connections.

$$C = \frac{\sum_{i \neq j} \max(h_{ik} + h_{jk})}{\sum_{i \neq j} \sqrt{(x_i + y_i) - (x_j - y_j)^2}}. \tag{2}$$

In the formula, h_{ik} and h_{jk} represent the number of hops from any position k nodes to anchor points i and j, respectively, and the maximum value of the sum of the number of hops from all position nodes to 2 anchor points is obtained and taken as the reference value of the number of hops between 2 anchor positions. This solution solves the estimation method of the average hop distance C, but the accuracy is equivalent to the classical DV-HOP algorithm, because the positioning accuracy is low, and locating the unknown node position will produce a large error rate.

$$\Delta S \leq \gamma \frac{\Delta L}{kL}. \tag{3}$$

The cloud computing approach has many distinctive features with the development of the times and is widely used in various fields, with several main forms: processing, backup, multi-point storage, and large-scale centralized storage. This computing approach has a high value of use, can be used across platforms, and can be configured with servers, databases and core software, and other resources to meet the actual needs of applications. Based on the above advantages of cloud computing technology, cloud
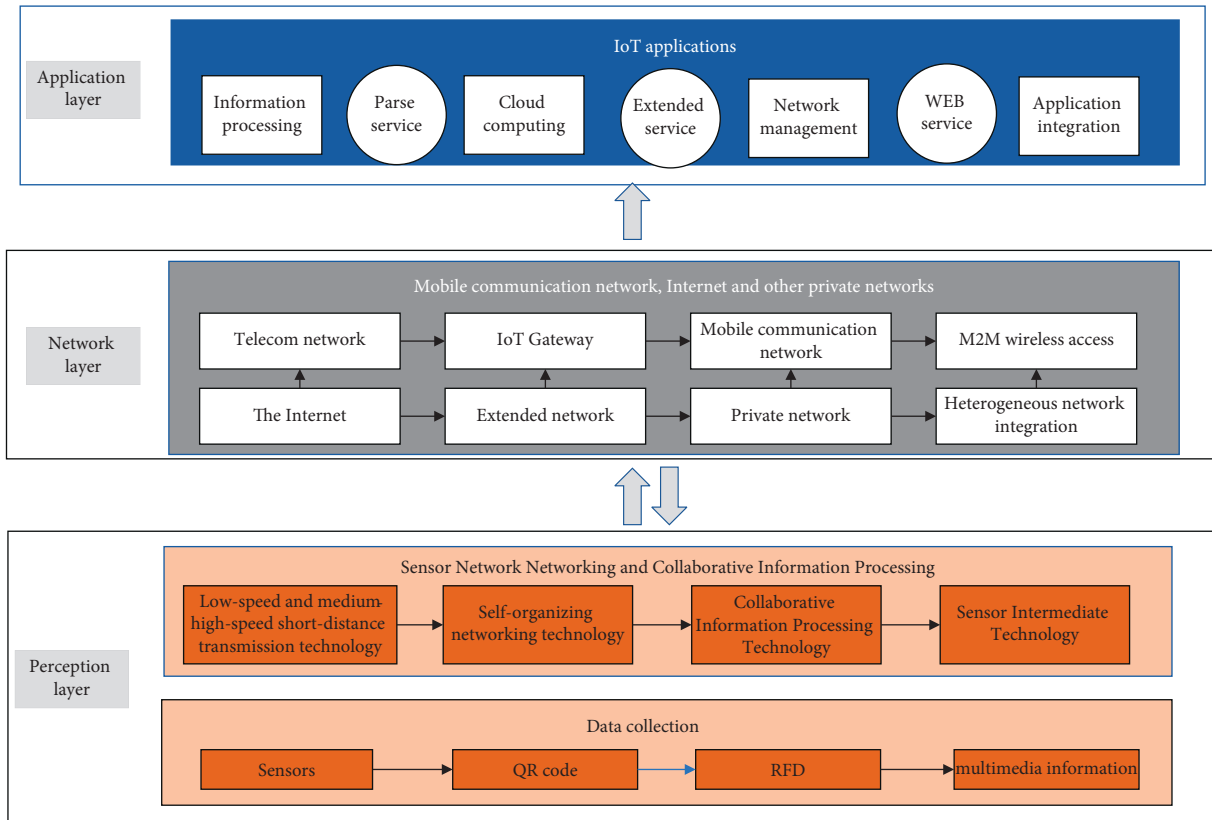
FIGURE 1: IoT network architecture diagram.
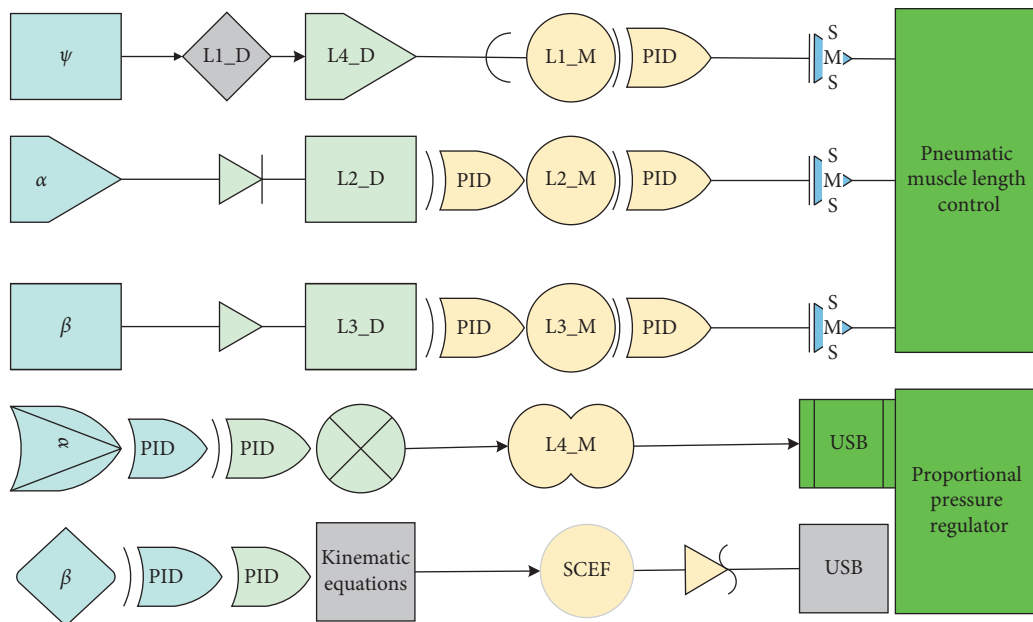
FIGURE 2: NB-IoT core network architecture diagram.

computing can provide stable and scalable basic services for storing and processing large-capacity WSN data. Contemporary society is the era of the Internet of Things; privacy issues are particularly important with the changing habits and learning paths of human life. From film, television, camera to video art, multimedia art, digital art, all are the

progress of science and technology that inspired the birth of new art. The progress of science and technology has created new platforms and spaces for artists, which has greatly enriched the forms of artistic creation and promoted the development of art. In a questionnaire on "What should we worry about with the development of the Internet of

Things," personal privacy is the most important concern for contemporary people, followed by security. In addition to these two reasons, the following issues are operating system, application vulnerability, unauthorized access, identity impersonation, etc. The detailed statistics are shown in Figure 3. For WSN with IoT-NB technology, the sensing nodes have a huge amount, high data redundancy, simplicity, weak computing power, security issues, and drawbacks in the case of relying only on cloud computing has a huge impact and threat.

3.2. Digital Media Content Creation Model Construction. Digital Media is not simply a superimposition of digital technology and traditional media, but full integration of digital media and communication content. Digital media is more commonly referred to as "New Media" in the discourse of Western researchers. The difference between "new media" and "old media" is that, on the one hand, the plural form of the word "media" means that "new media" is no longer seen as a single media organization or media structure (e.g., print media, radio, and television media, etc.), but as a collective concept that divides all forms of digital media into one [17]. On the other hand, the "new media" integrates the content and characteristics of all existing media forms and adds digital, interactive, two-way communication characteristics to them. The application of this new form of digital media and novel technology in art has a broad development prospect and market appeal. It is widely used in various exhibition halls, and digital media art is used in digital display, 3D virtual display, interactive display, and way-finding system design. Digital media can be seen as a system that covers digital media organizations, products, technologies, contents, networks, and terminals and is led by government departments, enterprises, and individuals, with the means of computer technology and network information technology; digital media art contents are disseminated through the Internet, mobile communication networks and digital TV networks, on computers, mobile personal terminals, wearable devices, etc. On the digital terminal, it presents various websites, multimedia, computer animation, electronic games, interactive novels, interactive images, interactive devices, and virtual reality products.

As information technology continues to evolve in the digital age, traditional media have subsequently taken on their digital new forms. Web-based communication methods are becoming increasingly popular, personal digital terminals are gradually entering people's lives, and the distance between digital media and human beings is gradually being brought closer. As Negroponte said in "Digital Survival," "the development of information technology will bring revolutionary changes to our way of life." Information will no longer be "sold" to "consumers," but humans will carry information provided by "digital handymen" and participate in the interaction process, and the mass media will gradually evolve into a two-way communication media with personal characteristics [18]. The mass media has gradually evolved into a two-way communication media with personalized characteristics.
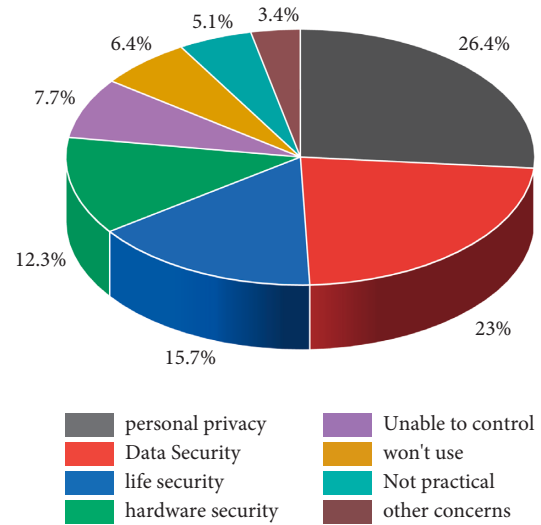
Figure 3: IoT security worry statistics results graph.

Chinese users have gradually shifted from traditional media to digital media and are spending more time on digital media than ever before. According to a statistic by Ariadne, from 2016 to 2021, the average daily use of digital media per capita in China increased from 1.78 hours to 3.08 hours, and the proportion increased from 35.8% to 50.4%, while the average daily exposure to traditional media per capita did not change significantly, as shown in Figure 4.

oT-NB can be called a simplified LTE network. It is characterized by wide coverage and large connections. It plays a very important role in the analysis and evaluation of IoT-NB IoT coverage and the planning of network capacity and scale effect. In the past, the exhibits on display were mainly in the form of physical objects. Due to their preciousness and state of conservation, these objects are not allowed to be touched by the public at close range, and there are restrictions on the time and place of their display. When the public goes to museums, they often encounter gaps in the content of the exhibits due to their condition or loans to other museums, which greatly affects visitors' interest and their desire to explore further in the exhibition activities [19]. With the addition of digital media technology, the exhibits start to develop from materialization to dematerialization and are presented in digital form in front of the visitors. Not only are the digital exhibits not easily damaged, is the display environment not restricted, and can the audience interact with them at zero distance, but also the same content can be displayed in multiple places at the same time, forming the sharing of resources. In addition, the digital exhibits will be a dynamic interpretation of static physical exhibits, making the exhibit content more vivid and comprehensive, stimulating the interest of the audience, thus improving the dissemination efficiency of museum display content.

A direct display scheme to solve the equations would make the algorithm less efficient, and therefore an acceleration of the algorithm is needed. The additive splitting algorithm (AOS) algorithm is a semi-implicit scheme for
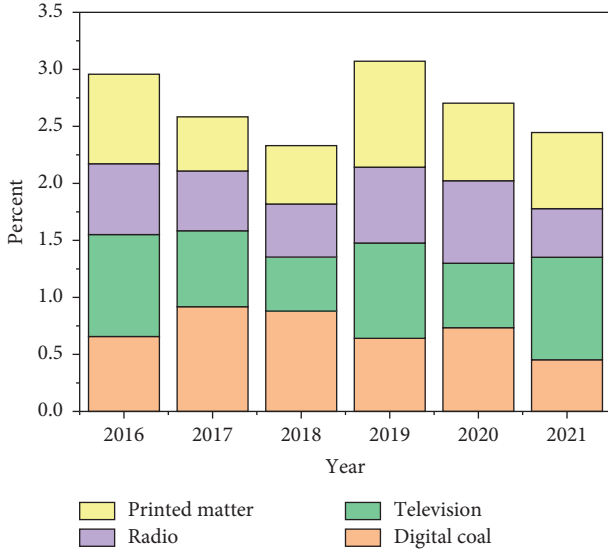
FIGURE 4: 2016–2021 average daily time and timeshare of people using each mainstream media.

solving the nonlinear diffusion equation, where the image data of the kth + 1st iteration is not directly generated from the data of the kth iteration but is given indirectly through a matrix-vector equation. Since the stability of this semi-implicit scheme is not affected by the iteration step, a larger time step can be set to obtain a faster update speed. The rows and columns of the image are first diffused in one dimension using the semi-implicit scheme.

$$I_{\kappa+1} = \sum_{d=1}^{2} \left(A_d + 2\tau I_\kappa \right) \cdot I_{\kappa+1},$$

$$C_l = \frac{2L_{\min}}{L_{\max} - L_{\min}}.$$

$$(4)$$

With the continuous progress of technology, people's information concept is also constantly updated. People hope to meet the needs of information acquisition through various dynamic image information display channels and have a variety of sensory experience. This is something that cannot be given by traditional exhibition activities. In the era of digital media technology, the popularity and development of digital media technology have overturned the form of media and formed a unique form of information dissemination, and the exhibition activities have also changed significantly with the intervention of digital media technology. The purpose of immersive experience is to make the audience be completely submerged in the virtual scene displayed by digital media technology [20]. It is revealed that media technology is of great significance to the change of human culture and life, and it has made an in-depth interpretation of media technology in various eras. It is believed that in the era of electronic media the time and space limitations will be broken, and the space and time distance between people will continue to shrink and form a "global village." If some parts of the space itself distract the viewer's attention, it means that the viewer is not fully engaged in the

virtual scene during the visit, which results in a weakened sense of psychological immersion. In the evaluation of the importance of each element, it can be found that the elements that highlight the image of the space itself are no longer considered important to enhance the experience of digital display, while the physical environment with low presence has become a new demand for visitors, as shown in the Figure 5.

Therefore, the principle of virtualization guides architects to weaken people's sense of place in physical space when designing space. On the one hand, they can achieve indirect blurring of material space at the visual level using special interface materials, environmental illumination, or environmental color treatment, so that people cannot judge the boundaries and environment of space in space, nor can they clarify orientation and direction. On the other hand, digital photoelectric technology can be used to create virtual boundaries, which can directly realize the materiality of some elements in the physical space from a technical perspective.

4. Analysis of Results

4.1. IoT Model Results. Wireless sensor network technology is a distributed sensor network technology that senses and checks external world sensors. Its communication is achieved mainly by wireless means, with a very flexible network setup that can be changed at will, as well as by using wireless or wired connections to the Internet. Wireless connectivity is a multi-hop organizational network model that combines tens of thousands of sensing nodes freely to form a network. The technology is characterized by its large scale, self-organization, and dynamic form, making it useful for a wide range of applications, such as medical, military, and environmental monitoring. However, the traditional wireless sensor network node deployment method has problems such as low coverage and long and unstable mobile steps. This study proposes a wireless sensor network node redeployment method based on frog hopping algorithm, simulating the way of frog flock foraging, classifying the subpopulation of network nodes, realizing the whole process of information transfer combined with global information exchange, mixing all the frogs within the subpopulation, then classifying the subpopulation, collecting a certain place in the network coverage, using the nearest node to collect relevant data information, reducing the repetitive collection situation occurs, calculating the deployment of wireless network nodes, within the scope of the network area, divided into small areas of equal size, where the nodes can cover the area, that is, the area being monitored, the need for special monitoring, to complete the wireless sensor network node redeployment. This algorithm is heuristic, although it can quickly converge to a more reasonable result, but because the node is the closest to the most correct solution, and the algorithm is relatively difficult and the cost is relatively high, there is no way to ensure a continuous extension of the network service life.

The coverage degree is an important parameter to measure the node deployment of the sensing network, i.e.,
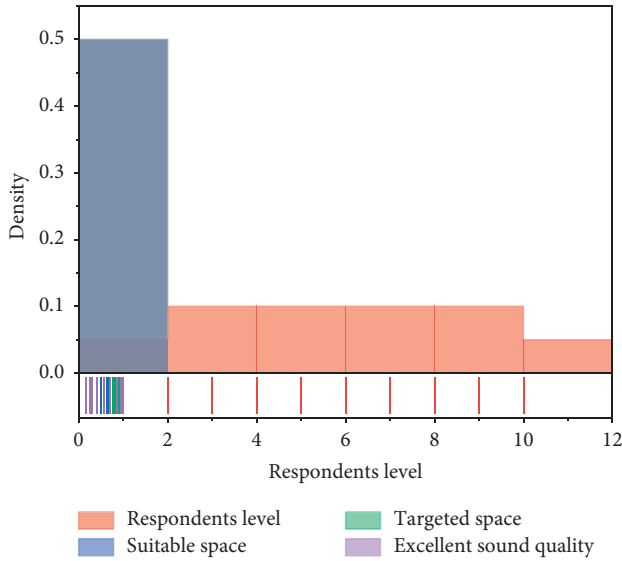
FIGURE 5: Respondents' level of need for a sense of presence in the space itself.

the area covered by all nodes and the total ratio of the target area. Therefore, node redeployment starts with the calculation of the coverage of the optimal solution node, and the node redeployment is achieved by searching the coverage area through the coverage rate. If a certain area is to be focused on monitoring within the detected area, the calculation method at this time is to be calculated using different weights. If only one of these areas is needed for special monitoring, this node coverage area is expressed by the following formula:

$$S_{\text{all}} = S_{\text{area}} - (w + 1) \times S_i. \tag{5}$$

The above equation represents the area of the key monitoring area and represents the weight, which can be changed according to different specialty areas. The random forest (RF) algorithm, first proposed by Tin Kam in 1995 and studied in more depth by Breiman, enables pattern classification and regression prediction of the data to be input. RF successfully implements a combination of bootstrap strategies and random subspace methods. Since the decision tree is built very fast, the training of the decision tree is much faster than training an artificial neural network. To implement the random forest-based greenhouse temperature prediction, three environmental factors, namely, air temperature, air humidity, and light intensity in the greenhouse, are selected as model inputs to build the random forest prediction model. In this paper, the number of random forest decision trees is 500, which helps to enhance the generalization ability of the model calculation process. The roles and functions of each structure of the Internet of Things technology are also different. The perception layer is mainly used for information perception, such as information collection, identification, etc.; the network layer is mainly used for the transmission of various information, and the application layer is mainly used for information. The 2000 sets of data are selected as the training set and 400 sets as the

prediction set. As shown in Figure 6, the comparison between the predicted output and the actual output shows that the predicted output is very close to the actual output, and the calculated RMSE is 0.0183, which shows that the prediction accuracy of the random forest model is good from the intuitive point of view.

From the above experimental results, the random forest prediction model has a low error in the prediction experiments of greenhouse temperature, which is a key factor for tomato growth, and its root mean square error of prediction is about 0.0183, which is a low prediction error and a more accurate prediction of greenhouse temperature trends. Coverage is a measure of test integrity, reflecting the energy consumed by the network coverage area, and the degree of node redundancy. The higher the coverage is, the less redundant the network nodes are, while the opposite is true for the redundancy of the network nodes.

During the process of recording the single-hop or multi-hop information sent at each position where all nodes stay, the minimum number of node hops is sent, that is, the shortest path. And the initial node communication radius R is 6 m. The simulation results show that the error rate is lower compared with the HOP-DV (random deployment) and the improved DVHOP-NBIOT (uniform deployment at the boundary) under the condition that the number of unknowns (12) is equal, the communication radius (6) is fixed, and the number of unknown nodes (31–13) is unequal. The accuracy of the unknown node localization is greatly improved. A comparison of the effect of different mobile anchor deployment methods on the error rate is shown in Figure 7.

The deployment method of mobile anchor locations has a direct impact on the localization accuracy of unknown nodes. The DVHOP-NBIOT optimized localization algorithm, from the perspective of the localization error, produces lower results using several deployment methods such as center uniform, square, and cross than the boundary average deployment method. The center-uniform deployment method gives the best results with an average error rate of 18.66%. In the HOP-DV classical algorithm, irregular location deployment always occurs, which leads to unstable and large differences in the localization accuracy between the anchor nodes to be located and the unknown nodes. The result of the final identification information is presented to humans, which further promotes the intelligent management of something by humans. In summary, the DVHOP-NBIOT optimization algorithm makes the best use of the self-localization capability of NB-IoT, the background computing capability, and further combines the deployment optimization of the original location of anchor points to calculate the unknown node positioning, and a comparative analysis based on the original DV-HOP method finds that the error rate is reduced to a great extent.

5. Digital Media Content Creation Simulation Experiment

The digital display is based on the theoretical knowledge of display design and digital theory, based on modern information technology and other related fields (such as

FIGURE 6: Schematic diagram of predicted and actual output distribution.

FIGURE 7: Comparison of the effect of different anchor deployment methods on the error rate.

activity of the audience, so visual is the main way for people to get information. On the one hand, the visual experience alone will affect people's sense of visiting experience and the degree of memory of information. Through the experiment, it was found that 94% of the information received came from visual, 11% from auditory, and 6% from the smell, touch, and taste perception, which indicates that visual and auditory are still the two main ways for people to obtain information, but the role of other senses also assists people in receiving information. In addition, the study also found that people's memory of information through reading accounted for 10% of the total, the memory of hearing for 20%, the memory of visual for 30%, the memory of listening while watching for 50%, the memory of telling by oneself for 70%, and memory of combining and acting on various ways for 90%, which indicates that the more perceptions involved in information acquisition, the stronger people's memory of information will be.

This section firstly introduces the basic concept of a variational method for image processing, the variational principle, and Euler's equation in detail and investigates the principle of a variational method for image restoration by introducing the variational problem in image restoration; i.e., the problem is transformed into a generalized function on the extremum problem to find the minimum value of the problem. The results show that the TV model has the strongest denoising ability, the P-M model has the strongest edge-holding ability, and the linear model is in the middle of the range, with a slightly weaker denoising ability than the TV model and weaker edge-holding than the P-M model. Finally, Chen's coupled PDE model is studied, which lays the foundation for multi-physics field coupling denoising. The system performance results are obtained by testing three sets of experimental data, as shown in Figure 8.

According to the consumption ratio, the consumption of lookup tables and flip-flops in the project is 38.02% and 14.89% respectively, both below 70%. The clock management MMCM and phase-locked loop PLL use 3 and 1, respectively, and there are still scalable parts of the clock. The global buffer BUFG uses 10 units, and the remaining 22 units can be used to reduce the signal transmission delay and increase the signal transmission frequency. In summary, the FPGA chip is fully capable of implementing this system project. Cloud computing methods mainly include processing, backup, multi-point storage, and large-scale centralized storage. This computing method has high use value, can be used across platforms, and can be configured with servers, databases, core software, and other resources to meet the actual needs of applications. As shown in Figure 9, due to the large error at the beginning, the single-neuron network adjusts the three PID parameters online quickly with a large amount of adjustment, so that it can compensate for the temperature error quickly and track the input signal as soon as possible; when the error becomes smaller, the single neuron adjusts the three PID parameters with a small amount, so that it can track the input signal stably and maintain the stability of the control.

This section presents the systematic implementation of the algorithm by modularizing the system and summarizing the design structure and design methodology of each module in the system in each subsection. Simulation of several

museums and exhibitions, etc.) to achieve the purpose of display for the public. Nowadays, digital image technology is maturing and has been popularized in China. Contemporary exhibition halls are also gradually adopting this new display technology, changing from static display to dynamic display, and a certain part of the halls even takes digital image technology to the extreme, developing into a multi-sensory interactive display mode. The growing maturity of digital technology has brought a revolution to the exhibition industry and brought a deeper impression and more vivid experience to the public. Digital display design will open a new chapter for this modern exhibition industry.

Traditional display activities are mainly in the form of static physical display, and browsing panels and exhibits are the main

FIGURE 8: System performance test results.

FIGURE 9: System parameter self-tuning curve.

module of the whole system, collects the sensors of each environmental factor at the terminal, sends the collected data to the remote server through the NB-IoT network, stores them in the database, and realizes the web page of B/S structure with Java technology. The administrator can log in to the web page through a PC or mobile terminal to realize intelligent management. In terms of control, the design adopts a fuzzy control algorithm to realize the regulation of external environmental factors and establishes the ARIMA model to predict the future environmental change trend and provide data support for intelligent management. There are still many shortcomings in the theoretical and practical research results in this paper. The topic of how to promote the development of the new media industry by promoting the development of digital media advertising is a grand and long-term topic that needs to be studied continuously. In the future research process, the author will conduct richer and more in-depth theoretical research to supplement and sublimate the research of this paper.

Data Availability

The data used to support the findings of this study are available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

Acknowledgments

References

[1] A. J. A. M. Van Deursen and K. Mossberger, "Any thing for anyone? A new digital divide in internet-of-things skills," *Policy & Internet*, vol. 10, no. 2, pp. 122–140, 2018.

[2] B. Jiang, J. Yang, and H. Xu, "Multimedia data throughput maximization in Internet-of-Things system based on optimization of cache-enabled UAV," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3525–3532, 2018.

[3] B. Al Hayani and H. Ilhan, "Image transmission over decode and forward based cooperative wireless multimedia sensor networks for Rayleigh fading channels in medical internet of things (MIoT) for remote health-care and health communication monitoring," *Journal of Medical Imaging And Health Informatics*, vol. 10, no. 1, pp. 160–168, 2020.

[4] A. A. A. El-Latif, B. Abd-El-Atty, and W. Mazurczyk, "Secure data encryption based on quantum walks for 5G Internet of Things scenario," *IEEE Transactions on Network and Service Management*, vol. 17, no. 1, pp. 118–131, 2020.

[5] S. Vadlamudi and D. L. Hargrove, "The Internet of Things (IoT) and social interaction: influence of source attribution and human specialization," *Engineering International*, vol. 9, no. 1, pp. 17–28, 2021.

[6] J. Qiu, Z. Tian, and C. Du, "A survey on access control in the age of internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4682–4696, 2020.

important modules, such as data storage and data processing modules, is also carried out and compared with the software implementation to highlight the advantages of the hardware implementation of the algorithm. Finally, the system is implemented on the board to complete the whole algorithm from design to implementation, which proves that the algorithm hardware implementation has the algorithm enhanced image function and meets the algorithm real-time requirements.

6. Conclusion

This paper designs a digital media content creation method based on NB-IoT IoT technology by studying the current stage of IoT technology, new energy technology, database technology, Web technology, and fuzzy control technology in the digital media field. The system abandons the traditional wireless communication technologies such as ZigBee and Wi-Fi, which are widely used, and chooses NB-IoT narrow-bandwidth IoT technology as the communication

[7] Y. Zhang, X. Ma, and J. Zhang, "Edge intelligence in the cognitive Internet of Things: improving sensitivity and interactivity," *IEEE Network*, vol. 33, no. 3, pp. 58–64, 2019.

[8] V. Mohammadi, A. M. Rahmani, and A. M. Darwesh, "Trust-based recommendation systems in Internet of Things: a systematic literature review," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, pp. 1–61, 2019.

[9] J. Chin, V. Callaghan, and S. B. Allouch, "The Internet-of-Things: reflections on the past, present and future from a user-centered and smart environment perspective," *Journal of Ambient Intelligence and Smart Environments*, vol. 11, no. 1, pp. 45–69, 2019.

[10] S. Arshad, M. A. Azam, and M. H. Rehmani, "Recent advances in information-centric networking-based Internet of Things (ICN-IoT)," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2128–2158, 2018.

[11] K. E. Jeon, J. She, and P. Soonsawad, "Ble beacons for internet of things applications: survey, challenges, and opportunities," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 811–828, 2018.

[12] F. Samie, L. Bauer, and J. Henkel, "From cloud down to things: an overview of machine learning in Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4921–4934, 2019.

[13] T. Qiu, N. Chen, and K. Li, "How can heterogeneous internet of things build our future: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2011–2027, 2018.

[14] C. K. M. Lee, Y. Lv, and K. K. H. Ng, "Design and application of Internet of things-based warehouse management system for smart logistics," *International Journal of Production Research*, vol. 56, no. 8, pp. 2753–2768, 2018.

[15] C. Puliafito, E. Mingozzi, and F. Longo, "Fog computing for the internet of things: a survey," *ACM Transactions on Internet Technology*, vol. 19, no. 2, pp. 1–41, 2019.

[16] A. Kumari, S. Tanwar, and S. Tyagi, "Verification and validation techniques for streaming big data analytics in internet of things environment," *IET Networks*, vol. 8, no. 3, pp. 155–163, 2019.

[17] E. Sisinni, A. Saifullah, and S. Han, "Industrial internet of things: challenges, opportunities, and directions[J]," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4724–4734, 2018.

[18] Y. Lu and L. Da Xu, "Internet of Things (IoT) cybersecurity research: a review of current research topics," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2103–2115, 2018.

[19] A. S. Wilner, "Cybersecurity and its discontents: artificial intelligence, the Internet of Things, and digital misinformation," *International Journal*, vol. 73, no. 2, pp. 308–316, 2018.

[20] A. T. Chatfield and C. G. Reddick, "A framework for Internet of Things-enabled smart government: a case of IoT cybersecurity policies and use cases in US federal government," *Government Information Quarterly*, vol. 36, no. 2, pp. 346–357, 2019.

Retraction

Retracted: English Speech Recognition System Model Based on Computer-Aided Function and Neural Network Algorithm

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Zhang, "English Speech Recognition System Model Based on Computer-Aided Function and Neural Network Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7846877, 11 pages, 2022.

Research Article

English Speech Recognition System Model Based on Computer-Aided Function and Neural Network Algorithm

Jin Zhang [ID]

School of Foreign Languages, Xinyang Agriculture and Forestry University, Xinyang 464000, Henan, China

Correspondence should be addressed to Jin Zhang; 2009280046@xyafu.edu.cn

With the economic globalization continuous growth of China's socioeconomic level tends to be internationalized, China's attention to English has been significantly improved. However, the domestic English teaching level is limited, so it is impossible to correct students' English pronunciation and make a reasonable evaluation at all times so that oral training has certain disadvantages. However, the computer-aided language learning system at home and abroad focuses on the practice of words and grammar, and the evaluation indicators are less and not comprehensive. In view of the complexity of English pronunciation changes, traditional speech recognition is difficult to recognize speech speed and improve its accuracy. Furthermore, to strengthen the English pronunciation of domestic students, a nonlinear network structure is studied in depth to simulate the human brain to analyze a model of speech recognition is established Mel frequency cepstrum characteristic parameters of human ear model and deep belief network. In this paper, the traditional computer pronunciation evaluation method is improved in an all-round way, and a set of high-quality speech recognition system of speech recognition method is constructed. Aiming at the above problems, it takes the students as the research, which proves that the method adopted in this paper can give the learners accurate pronunciation quality analysis report and guidance and correct their intonation and improve the learning effect, and the experimental data verify that the improved speech recognition system model recognition ability is higher than the traditional model.

1. Introduction

With the increase of globalization and China's internationalization, the Chinese people's demand for learning English is also easy to grow. However, because the Chinese vocal characteristics are different from those of the British, China does not have an English learning environment and lacks excellent English teachers. Due to the influence of time, location, and other factors, traditional classroom teaching can no longer meet the needs of English learning. At present, there are still great problems in English teaching; for example, teachers' descriptions are too general, and students' subjective initiative in learning cannot be fully brought into play, and learning English has become one of the research hotspots in the field of education. With the development of computer, the progress of language teaching methods, computer-aided language learning [1] provides the possibility to solve this problem. This technology allows learners to study in no time or area, can give learners timely and sharp evaluation and guidance, and can also help learners to find differences and standard pronunciation in different ways. Allowing learners to listen repeatedly, compare and correct pronunciation errors, and improve language learning efficiency.

The research goal of the call is to combine computer technology, especially speech recognition and evaluation technology, with the current language teaching methods to further establish a call system. Some famous international university research institutions are carrying out relevant research, such as Carnegie Mellon University in the United States, Tokyo University in Japan, and Cambridge University in the United Kingdom. MIT and Cambridge are working together to develop the oral communicative language learning project (SCILL), which aims to develop a conversational system to provide learners with conversational practice. WebGrader developed by Stanford University uses a speech recognition

engine [2] and scoring technology to score pronunciation exercises for second language learners. This conversational system allows students to practice spoken English online and promotes communication in English learning. In the research of speech quality evaluation in China, the Institute of automation, Tsinghua University, Microsoft Asia Research Institute, and other institutions are carrying out relevant content research and have achieved good results. The core of call is high-precision speech recognition is the important foundation of speech evaluation. DTW, HMM, and ANN are the classical mainstream methods, each with its advantages and disadvantages.

In recent years, the development of DL [3], BD [4], and CCT [5] and quality assessment technology have made rapid progress. DL originated from the field of ML [6], which aims to establish and simulate the deep neural network (DNN) of the human brain for analysis and learning. DNN can multilayer deep transmission of data by human brain neurons. Therefore, the study of English SRT based on DL can greatly improve the ability of voice aspect, improve access to information, and improve the user experience.

At present, with the rise of the international market, the training of spoken English in colleges and universities has been continuously strengthened. Therefore, the following problems are encountered in the study of spoken: Most people still choose language transponder and MP3 player, to learn oral English. However, none of these methods can carry out speech recognition, and they are limited to voice query and tracking functions. In addition, limited by technological factors, some call systems at home and abroad mainly concentrate on the learning of words and so on. Evaluation indicators are used as the evaluation, which can give an accurate score. Limited by their own level is not enough, and it is hard for them to find their mistakes and correct them with a total score of one point. In terms of oral test evaluation, the current oral test is still based on strong subjective will, different standards, and slow manual scoring. Due to the different experiences of scoring experts, and the different status of the same expert, the evaluation of the same voice will be biased. The English speech recognition system model constructed in this paper mainly uses the deep neural network to make the speech recognition can automatically recognize speech.

In order to solve those shortcomings, the DL application is selected, and a speech recognition model based on DBN is constructed [7]. Speech quality has been improved, and a reasonable speech quality evaluation model has been established considering the multiparameter indexes and their weights such as voice tone, voice speed, rhythm, and voice tone. The original achievement of the research can be applied to the study and training of human-computer interaction oral English. It can give learners more intuitive and useful evaluation, give advice, and give wrong pronunciation. It can quickly guide and motivate learners to learn English. The results are helpful to solve the problems in oral English teaching. In addition, English STR can improve the articulation of speech in the aspect of speech transmission, and the collection of causes is also more clear. The pronunciation of each syllable in English can be clearly collected, and the sound quality and timbre of speech can be guaranteed in the process of transmission.

2. Speech Signal Preprocessing and Feature Extraction

2.1. Basic Principles of SRT. SRT is one of the most popular text captions because its simple specification makes it easy to create and modify. The purpose of SR [8] is to decode the feature stream M of speech recognition and get the corresponding word sequence N, that is, to find $M = \text{argmax}_M p(N/M)$. It includes the following key modules: feature extraction, acoustic model, speech dictionary, language model. In general, it is difficult to search the optimal word sequence directly through features, so the transformation is based on Bayesian criteria.

$$P(M/N) = \frac{P(M/N)P(M)}{P(N)}, \quad (1)$$

where $p(M/N)$ is the generating acoustic feature M for a given word sequence N, which is the probability that the acoustic model needs to estimate. In the recognition model, if the speech features need to be recognized and extracted, the recognition program needs to be input into the system to complete the feature extraction. P (N) is the observed word sequence, which is the problem to be solved by the language model, and the construction of the speech recognition system is transformed into modular acoustic modeling, language modeling, word sequence decoding, and so on. The traditional speech recognition framework is shown in Figure 1.

(1) SSP and AFE: SSP [9] needs to preprocess and extract feature parameters. Generally speaking, the process of feature extraction is independent of the training of the recognition model and the testing of the system. According to the previous research experience, the feature parameter extraction of speech can be completed through a series of mathematical processes. The common speech feature parameters are linear prediction cepstrum coefficient and MFCC. MFCC parameters simulate the auditory characteristics of human ears. In most cases, MFCC is the most commonly used.

(2) Acoustic model [10]: HMM is a very popular acoustic model in speech recognition. The state jump model is consistent with the short-term stability of speech signals, so HMM can model continuous speech signals. In the GMM-HMM model, the probability of observation sequence is output by GMM, then the likelihood degree of observation sequence is obtained by HMM, and then, the optimal sequence path is obtained by decoding algorithm. With the continuous in-depth study of researchers, we also use DNN and HMM as acoustic models, in which DNN replaces GMM to generate the probability of observation traits, and HMM is used to describe the dynamic characteristics of speech signals. The model has been proved to be feasible and widely used in the speech recognition system.

(3) Language model [11]: M-Gram is most commonly used in speech recognition research. Language model based on RNN language model RNNLM is
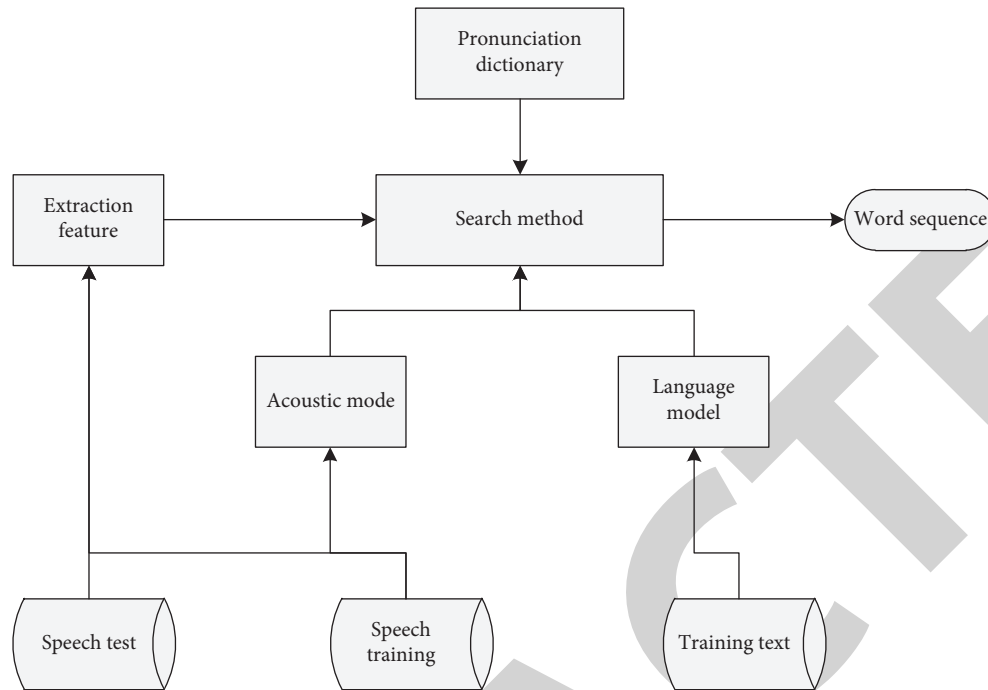
FIGURE 1: Structure of traditional speech recognition.

gradually introduced into speech recognition. This model can model long history information and has a better performance compared with n-gram. However, in large vocabulary speech recognition tasks, replacing n-gram with RNNLM will increase, and because of the calculation time, the recognition efficiency of the whole system will decrease.

(4) Decoding search: After the score of the acoustic model is obtained through the acoustic model, the language model is combined with the relevant algorithm to get the final recognition sequence. For example, the dimension bit algorithm is used to decode the HMM system, and some other models also use clustering search to obtain the optimal sequence.

2.2. Speech Signal Preprocessing.

Before speech signal pre-processing, it must be preweighted, framing, windowing, and endpoint detection. This is done to eliminate the influence of harmonics on the quality of speech signal, which is caused by human speech organs. SP affects the effect of speech, but a smoother speech signal can provide better parameters. Improving voice quality can solve coverage problems by reducing interference factors and improving the quality of the hardware.

2.3. Recognition of Speech Characteristic Parameters.

Speech recognition is the technology or function that enables a program or system to process human speech. It is mainly used to convert spoken language into computer text. The purpose of speech recognition is to make the machine finally understand the oral nature of human language and the form of a voice in the digital signal of the computer. It is obvious that the digital signal directly understood by the computer will make the whole model calculation huge. The voice signal is converted into the dimension that can reduce the characteristic parameters of speech and, at the same time, ensure the maximum retention of useful information. Speech features have better recognition characteristics, which is convenient for the modeling of speech recognition system and has the certain anti-interference ability to the environment. MFCC [12] is a cepstrum feature parameter extracted based on the critical frequency band effect of the human ear. In many researches, MFCC parameters and their differences are usually used to model speech recognition tasks.

Figure 2 describes the extraction structure of MFCC feature parameters. A complete speech is divided into equal length and adjacent frames by framing. In order to make the overlapped part of the adjacent frame signal smooth transition, Hamming window function is used to window the signal. Then, the fast Fourier transform is used to transform. Then, the Mel triangle filter is used to calculate the signal, and then, the logarithm is taken to get the Mel spectrum of the signal. Finally, MFCC parameters are obtained by DCT. MFCC parameters reflect the static characteristics of speech. Usually, after getting the parameters of MFCC, the dynamic change of speech is described by calculating the difference, and the recognition performance of the system is improved by combining the static and dynamic methods. For example, MFCC can obtain the time length of speech and then divide it into N speech segments for recognition according to the length of the speech.

3. Research on Speech Recognition Model

3.1. Concept of NN.

NN [13] can be regarded as bionic because the neural network is designed according to the
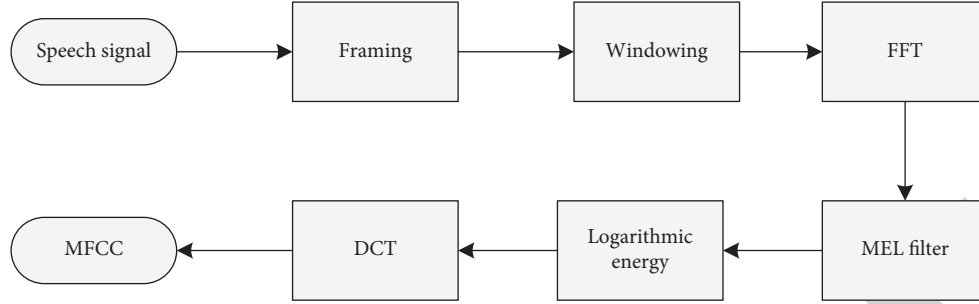
Figure 2: MFCC parameter extraction process.

human brain nerve and have the function of calculation and decision-making. The neural network is mainly used in classifiers. The NN calculates layer by layer according to the input characteristic quantity and finally outputs the identified signal. The NN itself has the function of learning. In the training stage, it can learn the unique characteristics of different samples according to the training samples and carry out memory; in the test samples, when encountering similar samples, it can distinguish the type of samples according to "memory," so the NN needs a large number of complete samples, and can correctly identify new samples through continuous learning and memory. It can ameliorate the generalization of NN. In the process of learning, if the training samples are not enough, it will lead to underfitting, and the actual classification effect will be very poor. If the training samples are single or overtrained, there will be overfitting. The overfitting network has a high recognition rate in training samples, excellent performance, and low recognition rate in test samples.

Different types of neural networks have different application scenarios. Common neural networks include BP NN [14], convolution neural network [15], wavelet NN [16], radial basis network [17], and self-organizing NN [18], among which BP NN is mainly for classification; convolution NN is mainly used for image recognition. Although the neural network is widely used in engineering, such as fingerprint recognition, face recognition, and biological body recognition, it is still unable to discuss the specific learning logic in theory, such as how to learn and remember the weight or value according to a certain function change and how to improve the number of hidden layers of the NN. In engineering, the parameters of the neural network are mainly adjusted, and the best is determined by the method according to the results of identification. With the increase of engineering complexity, there will be more and more related parameters, and the task of related parameter adjustment will gradually increase.

The artificial neuron [19] can be expressed by the following formula:

$$\begin{cases} a = \sum_{i=1}^{r} \omega_i x_i \\ b = f(a + \theta) \end{cases} . \tag{2}$$

Here, $x_1, x_2, \ldots, x_r$ represents the r input of the neuron; ω_i represents the connection strength (connection weight) of the ith input; θ is the bias of the neuron; b is the neuron. Therefore, the artificial neuron is a multi-input single-output nonlinear structure.

3.2. BP Neural Network.
Artificial NN, also known as NN, is a data processing model based on biological NN to each other for calculation and changes its structure according to external information, mainly through adjusting the weight of the network training input data and model, so that it has the ability to finally solve practical problems.

At present, a variety of neural network models have been put forward in the academic field, including radial basis function network [20], Hopfield network [21], CMAC cerebellum model [22], and BP NN. This paper studies BP NN. BP NN is usually multilayer, and another related concept is MLP, which has multiple hidden layers. MLP emphasizes the structure of NN, while BP NN uses the BP algorithm to adjust the network weight on the premise of a multilayer network. In most cases, they usually refer to the same network.

3.3. Deep Learning NN.
Deep NN is defined as a perceptron (MLP) with multiple layers that is usually greater than 2. Figure 3 depicts a NN with one input layer, one output layer, and three hidden layers. Neurons between adjacent layers are connected by weight. Through the input v_1, weight W_1, and deviation b_1 of this layer, the input of the next layer can be obtained by using the activation function.

$$v^{l+1} = f(v^l W^l + b^1). \tag{3}$$

Error backpropagation (BP) method is the deep NN. By calculating the partial derivative of each weight and the deviation of the loss function, the descent gradient of each parameter is obtained, and then, the parameter is updated by the gradient.

The depth NN can be represented in many different ways by different structures of hidden layer elements and different connection ways between adjacent layers. In speech recognition tasks, RNN and its variant CBLSTM have achieved good results and have been widely used.

The RNN connects the hidden layer elements of the feedforward neural network. When calculating the output of the hidden layer node, its calculation result depends on the calculation result of the previous node. Figure 3 depicts the RNN structure:
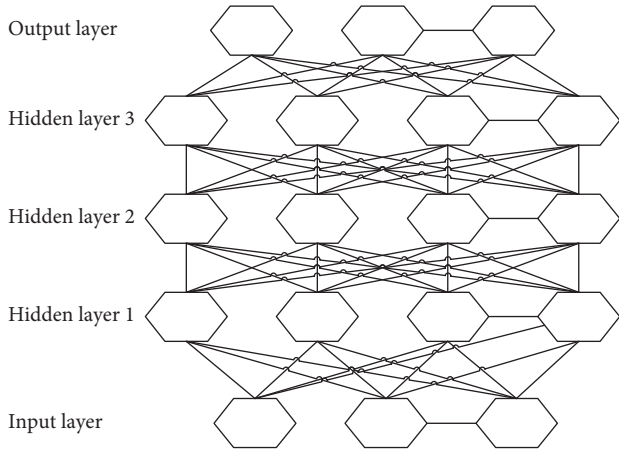
Figure 3: Depth neural network model.

The hidden layer of RNN is self-connected to expand the hidden layer. U, V, and W are connection output layer and hidden layer, respectively. The connection weights of all nodes are shared. x_t is the input of the current node, that is, the characteristic parameter vector obtained from the current voice frame signal, and the calculation formula of the hidden layer s_t and the output o_t are as follows:

$$s_t = \tan h(x_t U + x_{t-1} W + b_s),$$
$$o_t = s_t V. \tag{4}$$

The advantage of RNN hidden layer self-connection is that by extending it, the network of any length in the sequence can be obtained. In addition, different speech lengths are inconsistent. It provides a very flexible input model for speech signal by using the hidden layer self-connection of RNN.

3.4. Research on Speech Recognition Based on DBN. The core of the deep belief network is a layer greedy learning algorithm and then uses the backpropagation algorithm to fine-tune the network to get a better performance network model. Experiments show that the weight initialization of the multilayer perceptron corresponds to this network model.

Deep Boltzmann machines (DBMs) can be obtained by using an unsupervised greedy layer-by-layer approach. As shown in Figure 4 (i.e., the Bayesian belief network does not connect the layers between nodes) to approach the visual layer, layer by layer error propagation, and contain the farthest part used from the visual layer, we can obtain the deep belief network speech recognition model (DBN). In essence, RBM training can obtain the global optimal initial parameters, so as to improve the network performance [23, 24].

3.5. Content of Speech Recognition. As shown in Figure 5 first of all, the voice analog signal is digitized, and the voice signal is collected by the sound card of PC. According to Nyquist sampling theorem, in the process of analog/digital signal conversion, when the sampling frequency f_{s_max} is greater than twice the highest frequency of F_{max} in the signal, the following formula is obtained:

$$f_{s\_max} \geq 2 * F_{max}. \tag{5}$$

The sampled digital signal can fully express the information in the speech. The normal speech is generally 50~5000 Hz, and this paper sets the sampling as B kHz including preweighting and window. Preprocessed speech signal was collected. The speech features that can be recognized by the recognition model include auditory features, because the ultimate purpose of speech is to be listened to by human beings, and speech also has specific timbre features. Finally, the selected speech feature is recognized by the speech recognition model designed in this paper.

4. Evaluation of Multiparameter Pronunciation Quality

4.1. Process for Assessing Pronunciation Quality. The judge of voice quality is divided into subjective evaluation and objective evaluation, as shown in Figure 6.

Objective evaluation is the prediction and estimation of subjective evaluation, so the accuracy of subjective evaluation directly affects the performance of objective evaluation. Only when the experts evaluate the speech quality of the test speech truly and reliably and, on this basis, carry out machine evaluation, the objective evaluation is more meaningful. In order to get a real and reliable expert score, it is necessary to establish an appropriate voice quality evaluation standard according to the application requirements.

According to different situations, this paper uses two objective scoring methods to calculate scores.

(1) Quadratic curve fitting method is mainly used for the verification of single speech feature experiment. The specific process is as follows: the target distance between the training samples and the reference speech is calculated by using the training samples in the language library, and then, the parameters of the quadratic fitting curve are calculated by using the curve fitting method according to the principle of minimum variance. According to the parameters of the quadratic fitting curve, the subjective and objective mapping test is carried out, and the corresponding relationship between the objective distortion measurement $D(k)$ and the subjective estimation MOS is obtained.

$$M = a \cdot D^2(k) + b \cdot D(k) + C \quad 1 < k \leq K, \tag{6}$$

where *a, b, C* are the parameters of the quadratic fitting curve obtained by training, M is the target distance value, $D(k)$ is the objective estimate of MOS subjective score obtained by the quadratic fitting curve, and K is the total number of distortion conditions.

(2) BP scoring method mainly aims at the overall evaluation system of this paper, that is, using the BP algorithm to calculate the objective score of test speech.
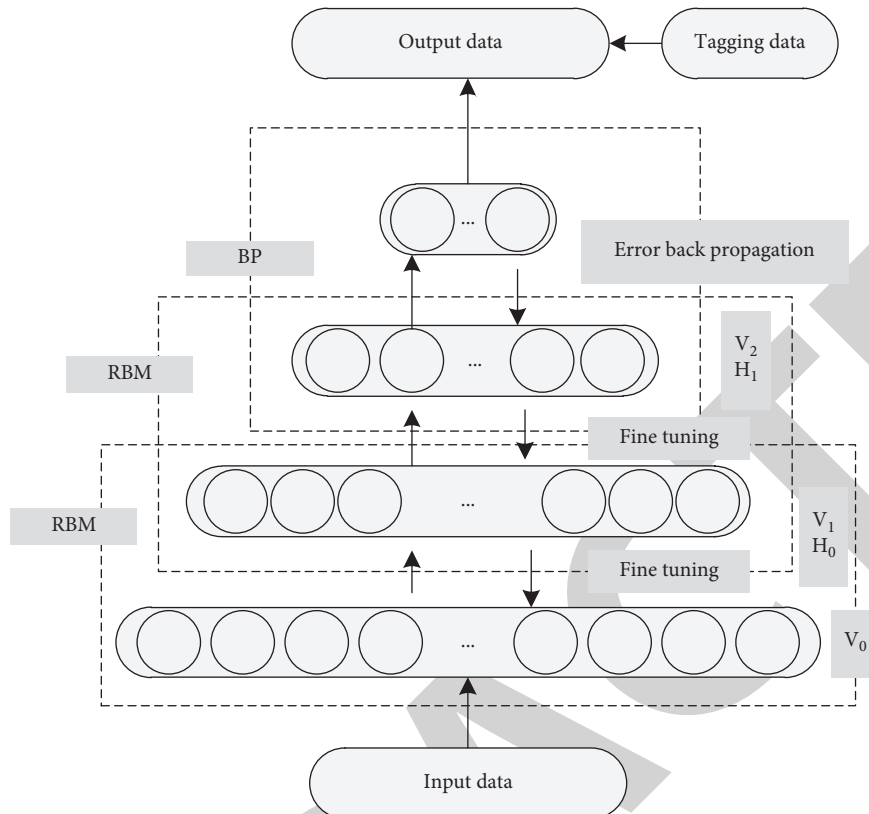
Figure 4: DBN voice model.
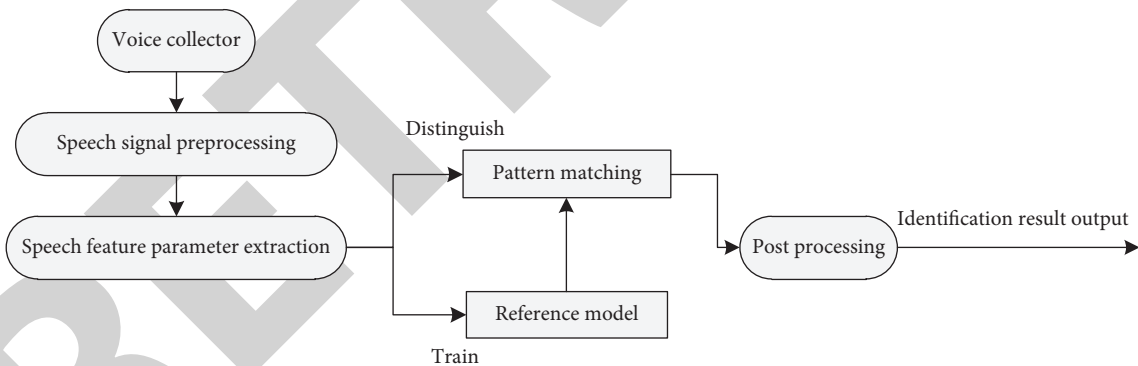
Figure 5: Content of speech recognition.

4.2. Quality Evaluation Index. In recent years, Chinese English learners have made a great breakthrough in the overall level of pronunciation, especially in the pronunciation of monosyllabics (vowel-consonant), such as vowels in the back and change. This is largely due to the country's investment in educational resources and the enthusiasm of English learners. Therefore, for most English learners, they can better master the pronunciation of English words. But in real life, the rhythm of sentences is the key to English fluency. Many learners encounter a bottleneck after mastering the pronunciation of a single sound because they find that they can pronounce each sound very standard but still can express English with the strong Chinese flavor. This group

includes not only beginners but also many intermediate-level learners, English majors, and even some people who use English for a long time. In a sentence, prosody is a very important factor. Every language has its own prosody.

The evaluation of voice quality mainly includes voice tone, voice length, and voice rhythm. In the process of evaluation, phonemes and words are mainly evaluated by intonation. Sentences and paragraphs should not only check the content of the pronunciation itself but also determine the true meaning of the sentence to a large extent. When evaluating the pronunciation quality of sentences and paragraphs, we need to consider the prosody information comprehensively, such as whether the speaker can grasp the
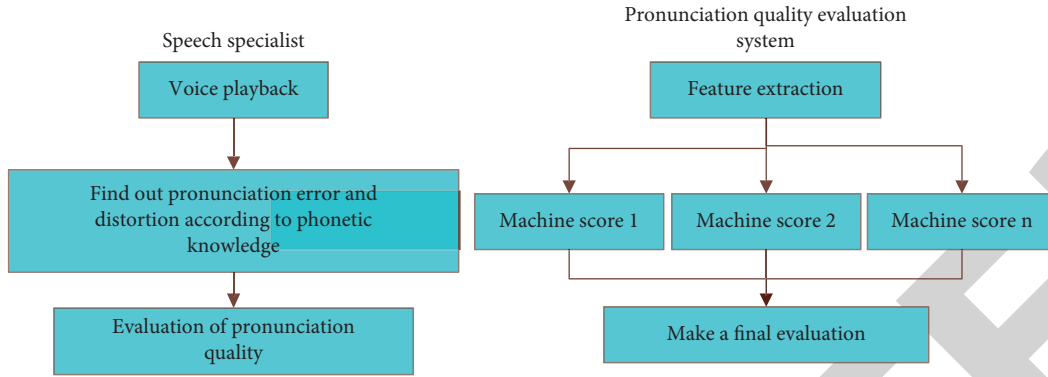
FIGURE 6: A comparison between subjective and objective assessment of pronunciation quality.

key information of the sentence more accurately, whether the sentence is not so important, and whether the length of the voice is appropriate. In other words, when evaluating the voice quality of English sentences, good voice quality not only requires complete, no pronunciation errors but also needs speaking speed, accent, and accuracy. Therefore, this paper uses intonation and prosody to evaluate the quality of speech.

From the perspective of linguistics, prosody, also known as rhythm, refers to the organization of phonemes higher than the level of segments. It is a systematic organization that organizes various language units into discourse or related chunks in discourse. The realization of prosody can not only convey linguistic information but also paralinguistic and nonverbal information (including emphasis, emphasis, phrase grouping), as well as emotion, attitude, vocabulary discrimination, and other functions. Therefore, the quality of English pronunciation is closely related to its rhythm.

Prosodic features, also known as suprasegmental features, mainly refer to speed, rhythm, and intonation composed of dynamic modes such as pitch, intensity, and duration. This paper uses speech speed evaluation and paired variation basic frequency intonation to evaluate and guide English speech quality.

4.3. Evaluation Model of Multiparameter Pronunciation Quality. Different groups (such as college students and business people) have different requirements for learning oral English. Based on the evaluation of voice tone, speed, rhythm, intonation, and other quality indicators achieved by taking oral English of college students as the research object, a comprehensive analysis of various indicators is made. In the comprehensive quality evaluation, the relationship between the weight of each index should be considered, and a multiparameter evaluation model and method of college students' English pronunciation quality should be established to evaluate the pronunciation quality reasonably and objectively.

4.4. Speech Evaluation Experiment. The purpose of the speech evaluation experiment is to verify the performance of the English speech quality evaluation model and method

proposed in this paper. In this paper, consistency rate and PCC are used to express the manual evaluation.

The consistency ratio refers to the ratio of sample size, which is consistent with machine evaluation and nonautomatic evaluation. The specific calculation method is as follows:

$$A_{\text{Same rate}} = \frac{\text{Evaluate the same number of samples}}{\text{Total sample size}}. \quad (7)$$

The adjacent consistency rate is the ratio between the manual evaluation and the sum of adjacent samples and total samples, in which "adjacent" is defined as one level difference between machine evaluation and manual evaluation. The specific calculation method is as follows:

$$A_{\text{Adjacent consistency rate}} = \frac{\text{Same number of samples + adjacent samples}}{\text{Total sample size}}. \quad (8)$$

The two variables, which is expressed by r, are the linear correlation between two variables, and the value range is $-1 \sim +1$. The greater the absolute value, the stronger the correlation. Generally, $0 < R < 0.2$ indicates very weak correlation, $0.2 < R < 0.4$ indicates very weak correlation, $0.4 < R < 0.6$ indicates medium correlation, $0.6 < R < 0.8$ indicates very strong correlation, and $0.8 < R < 1$ indicates very strong correlation. The specific calculation method is as follows:

$$r = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - \left(\sum x_i\right)^2} \sqrt{N \sum y_i^2 - \left(\sum y_i\right)^2}}. \quad (9)$$

5. Experimental Simulation Results and Analysis

5.1. Reliability Analysis of Human-Computer Evaluation of Pronunciation Quality. It can be seen from Figure 7 that the deviation of a sample number of manual rating and machine rating of the test set in each score segment is kept in a reasonable range. As a whole, this distribution pattern of rating sample conforms to the performance distribution characteristics of general tests, and from the data distribution and data amount, the machine rating meets

the general requirements of test rating. However, from the micro point of view, manual rating and machine rating still show differences. The score of manual rating is relatively high, and the machine rating will be "stricter."

5.2. Analysis of Speech Recognition Experiment Results.
Compared with that of other models, this paper uses the Spanish Arabic digital, which includes 8800 Arabic digital voice data (88 people pronounce 10 Arabic digital voice, and each digit repeats 10 times). 6600 pronunciations of the first 66 subjects were used as training sets, and 2200 pronunciations of the last 22 subjects were used as test sets. The experimental software is matlab82013a.

For the same UCI machine learning knowledge base, Arabic spoken digital data set, Nacereddine Hammami, a new model of pattern tree distribution (TDA–GTS) based on tree structure and a new model of pattern tree distribution (TDA–MWST) based on regenerating tree structure are proposed. Compared with the traditional CDHMM and the CDHMM, the recognition effect is improved. Huang Wentao proposed a minimum KASWT weight value, and compared with the BP_Adaboost algorithm, its recognition effect is significantly improved.

As can be seen from Figure 8, the recognition rate of the DBN model is 97.32%, which is better than the above model. Therefore, the DBN established in this paper is reasonable and effective, which can be further used for speech quality evaluation.

5.3. Artificial Evaluation Experiment.
In order to facilitate the calculation, according to the method described, after obtaining the scores of 240 statements out of 10 sentences of 24 students, the scores are compared with manual evaluation as shown in Table 1 and Figure 9.

A total of 210 samples were obtained, including 34 first-level samples, 2 second-level samples, and 0 third-level samples. The results show that the coincidence rate of machine and artificial pitch is 87.63%, and the adjacent coincidence rate is 99.79%.

In the aspect of speech rate evaluation, there are 185 samples of ME and manual evaluation, 48 samples of the first grade not same, and no samples of the second- or third-grade difference; it shows that the consistency rate of machine and manual speech rate evaluation is 83.21%, the adjacent consistency rate is as high as 99.99%, and the Pearson coefficient is 0.527, which shows that the speech rate EM in this paper is credible.

In the aspect of result, there are 212 samples in the same level, 31 samples in the first level, 4 samples in the second level, and no difference in the third level. The results show that the consistency rate of machine and artificial RE is 86.89%, the adjacent consistency rate is 98.63%, and the Pearson coefficient is 0.556.

In terms of intonation, there are 184 samples with the same level of ME and manual evaluation, 39 samples with the first level are not the same, only 3 samples with the second level are not the same, and no samples with the third level are not the same. It shows that the consistency rate of
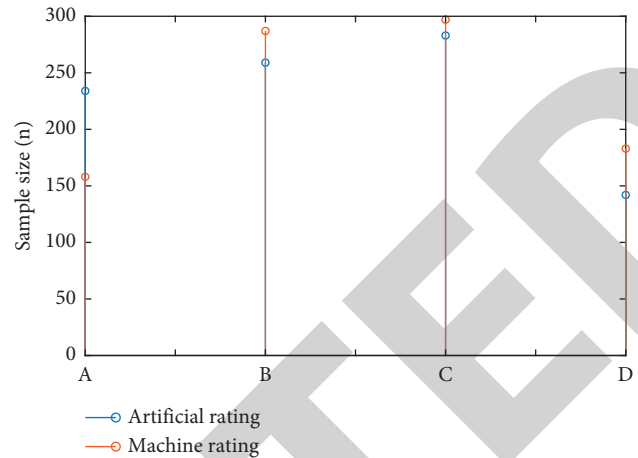
Figure 7: Distribution chart of human-machine rating.
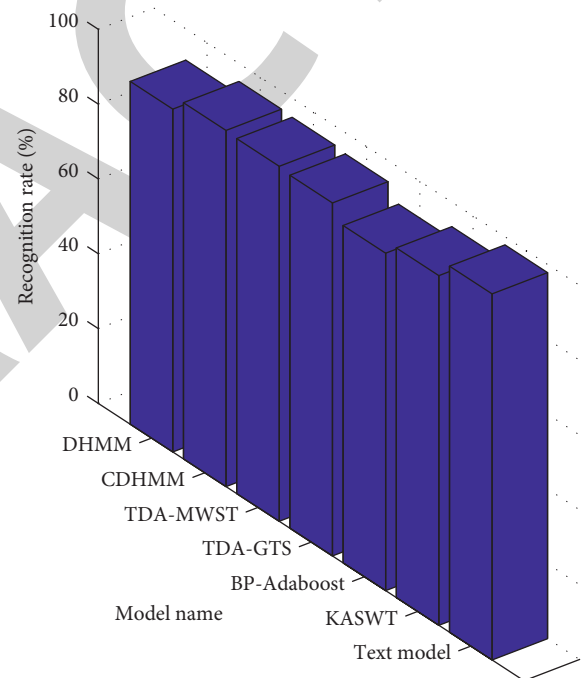
Figure 8: Comparison of recognition rates of different methods.

machine and manual intonation evaluation is 81.36%, and the adjacent consistency rate is as high as 99.12%. The results show that the intonation evaluation method is very important.

To sum up, the assessment methods of intonation, rhythm, and intonation adopted in this paper are reliable and can be further applied to the construction of the English speech quality assessment model.

5.4. Test Results of Evaluation Model.
This paper evaluates 250 sentences out of 11 sentences of 25 students. The experimental results are shown in Figure 10. There are 200 samples for machine evaluation and manual evaluation of the same level, 28 samples for level 1, and no

TABLE 1: Sample number of experimental results of evaluation index.

Index/sample number	Identical	Difference grade 1	Difference grade 2	Difference grade 3
Intonation	214	35	2	0
Speech rate	189	36	1	0
Rhythm	211	32	2	0
Intonation	187	33	3	0
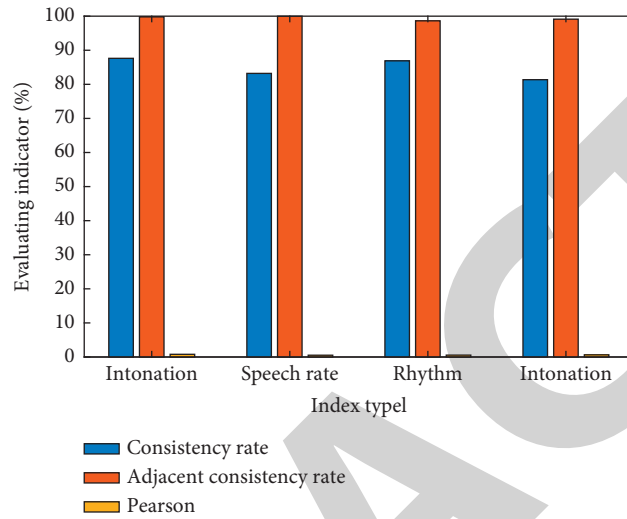
FIGURE 9: Evaluation index experimental results, statistical index.
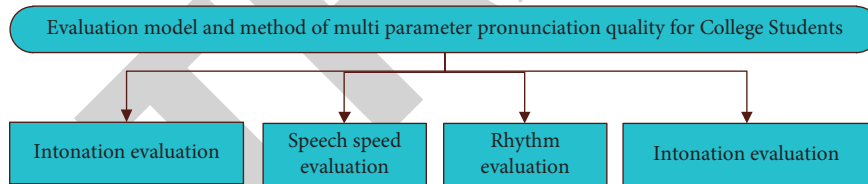
FIGURE 10: Evaluation model of multiparameter pronunciation quality for college students.

TABLE 2: Overall evaluation test results.

Index/sample number	Consistency rate (%)	Adjacent consistency rate (%)	Pearson
Intonation	88.98	100	0.833

difference between level 2 and level 3 (see Table 2) for specific data.

According to Table 2, intonation is the most important, followed by rhythm and intonation, and speed is the least important. After obtaining the above experimental results, we communicated with the pronunciation experts to further discuss the reliability of the proposed multiparameter speech quality evaluation model. Experts believe that intonation is the most important index to evaluate the quality of English pronunciation, which requires accurate content, and pronunciation errors mainly show the emotional color of the speaker, enhance the melody of the voice, and make the voice closer to the reality of life.

6. Conclusion

Based on the globalization of China's economy, English has become a popular language in China. However, due to the Chinese accent, English is regarded as "Chinglish." However, the traditional speech recognition algorithms DTW, HMM, and ANN are not comprehensive enough, because these speech recognition models ignore that speech quality will be damaged during transmission. Because most of the computer-aided language learning systems at home and abroad focus on the accumulation of grammar and words, the importance of pronunciation is ignored, lack of evaluation indicators in pronunciation as the basis for evaluation,

so as to better correct students' pronunciation and improve oral skills. In addition, the traditional oral test has a strong subjective will, which is only based on manual scoring, so it cannot be fair. In view of the various problems, this paper strengthens the traditional speech recognition technology to consider many factors such as intonation, intonation, speed, and rhythm and further refines the defects of traditional speech recognition technology. From the tables in this paper, it can be clearly concluded that multiparameter pronunciation quality evaluation can be considered, which can give students accurate, effective, and objective evaluation and guidance, improve learners' oral English learning effect, and achieve a qualitative leap [25].

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Muljono, U. Afini, C. Supriyanto, and R. A. Nugroho, "The development of Indonesian pos tagging system for computer-aided independent language learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 12, no. 11, p. 138, 2017.

[2] P. Zhang, Z. Ji, W. Hou, X. Jin, and W. Han, "Design and optimization of a low resource speech recognition system," *Journal of Tsinghua University*, vol. 57, no. 2, pp. 147–152, 2017.

[3] H. Ye, G. Y. Li, and B.-H. F. Juang, "Power of deep learning for channel estimation and signal detection in ofdm systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2017.

[4] Z. Zhang, W. Zhao, J. Xiao, Y. Wang, and M. Sun, "The big data center: from deposition to integration to translation," *Nucleic Acids Research*, vol. 45, no. D1, pp. D18–D24, 2017.

[5] A. Cardoso, F. Moreira, and D. F. Escudero, "Information technology infrastructure library and the migration to cloud computing," *Universal Access in the Information Society*, vol. 17, no. 1, pp. 1–13, 2017.

[6] F. Cabitza, R. Rasoini, and G. F. Gensini, "Unintended consequences of machine learning in medicine," *JAMA*, vol. 318, no. 6, p. 517, 2017.

[7] J. Zheng, X. Fu, and G. Zhang, "Research on exchange rate forecasting based on deep belief network," *Neural Computing & Applications*, vol. 31, no. 1, pp. 573–582, 2019.

[8] C. Kim and R. M. Stern, "Power-normalized cepstral coefficients (pncc) for robust speech recognition," *IEEE/ACM Transactions on Audio Speech & Language Processing*, vol. 24, no. 7, pp. 1315–1329, 2017.

[9] M. G. Blevins, S. L. Bunkley, E. T. Nykaza, A. Netchaev, and G. Ochi, "Improved feature extraction for environmental acoustic classification," *Journal of the Acoustical Society of America*, vol. 141, no. 5, p. 3964, 2017.

[10] M. Naderi, A. G. Zajic, and M. U. Patzold, "A non-isovelocity geometry-based underwater acoustic channel model," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 2864–2869, 2017.

[11] E. P. Frigieri, T. G. Brito, C. A. Ynoguti, A. P. Paiva, and P. P. Balestrassi, "Pattern recognition in audible sound energy emissions of aisi 52100 hardened steel turning: A mfcc-based approach," *International Journal of Advanced Manufacturing Technology*, vol. 88, no. 5, pp. 1383–1392, 2017.

[12] W. He, Y. Chen, and Y. Zhao, "Adaptive neural network control of an uncertain robot with full-state constraints," *IEEE Transactions on Cybernetics*, vol. 46, no. 3, pp. 620–629, 2017.

[13] B. Wang, X. Gu, L. Ma, and S. Yan, "Temperature error correction based on bp neural network in meteorological wireless sensor network," *International Journal of Sensor Networks*, vol. 23, no. 4, p. 265, 2017.

[14] M. M. Khan, A. Mendes, and S. K. Chalup, "Evolutionary Wavelet Neural Network ensembles for breast cancer and Parkinson's disease prediction," *Plos One*, vol. 13, no. 2, Article ID e0192192, 2018.

[15] T. Li, S. Duan, J. Liu, L. Wang, and T. Huang, "A spintronic memristor-based neural network with radial basis function for robotic manipulator control implementation," *IEEE Transactions on Systems Man & Cybernetics Systems*, vol. 46, no. 4, pp. 582–588, 2017.

[16] B. Serrien, M. Goossens, and J.-P. Baeyens, "Issues in using self-organizing maps in human movement and sport science," *International Journal of Computer Science in Sport*, vol. 16, no. 1, pp. 1–17, 2017.

[17] X. Zhang, W. Wang, Q. Liu et al., "An artificial neuron based on a threshold switching memristor," *IEEE Electron Device Letters*, vol. 39, no. 2, pp. 308–311, 2018.

[18] M. A. B. Mansor, M. S. B. M. Kasihmuddin, and S. Sathasivam, "Robust artificial immune system in the hopfield network for maximum k-satisfiability," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 4, p. 63, 2017.

[19] J. Chen, P. Li, G. Song, and Z. Ren, "Control of an innovative super-capacitor-powered shape-memory-alloy actuated accumulator for blowout preventer," *Modern Physics Letters B*, vol. 31, no. 1, Article ID 1650426, 2017.

[20] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, and K.-R. Muller, "Evaluating the visualization of what a deep neural network has learned," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 11, pp. 2660–2673, 2017.

[21] F. Shuang and C. L. Philip Chen, "A fuzzy restricted Boltzmann machine: Novel learning algorithms based on crisp possibilistic mean value of fuzzy numbers," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 1, pp. 117–130, 2018.

[22] M. Tohtayong, S. Khan, M. Yaacob et al., "The combination of Newton-raphson method and curve-fitting method for pwm-

Research Article

Evaluation Method of Financial Accounting Quality in Colleges and Universities Based on Dynamic Neuron Model

Lu Liu [ID]

Shandong Sport University, Jinan, Shandong 250102, China

Correspondence should be addressed to Lu Liu; liulu@sdpei.edu.cn

With the deepening of reform and opening up, great changes have taken place in the university financial management system. The role of financial analysis in university activities is becoming more and more obvious. In the new environment, especially in university financial reporting, we must establish effective reasonable and scientific financial analysis index system and quality evaluation team. In order to reflect the financial situation of colleges and universities, the university financial analysis indicators in this field have important theoretical and practical significance, such as finance, budget implementation, effective utilization of funds, risk prevention, and the formulation and application of such indicators. The financial management level of colleges and universities is improved, and the scientific development of colleges and universities is promoted. In this paper, we introduce the dynamic model of neurons, design a learning algorithm, and apply it to the quality evaluation of financial reports in colleges and universities. Through this research, a single-layer feedback network capable of fast learning and learning is established. This is not only helpful for universities to evaluate the quality of financial accounting business. However, enriching the significance of financial management in higher education has theoretical value.

1. Introduction

Financial analysis of colleges and universities is an important means to improve the efficiency of capital use and management level, an important indicator to promote the school to improve the financial risk tolerance, and an important basis for the school leaders to understand the current situation and make decisions for the future, so it has put forward higher and more requirements for the financial analysis of college system, and it has become vital to actively explore the way to optimize the financial analysis of colleges and universities [1–3]. The connotation of financial analysis of college system refers to the systematic analysis, comparison, and research of business activities of colleges and universities within a certain period of time by using special methods based on and starting from business plans, accounting statements, and other relevant information and summarizing and evaluating the experience of financial management, revealing the existing problems, gradually understanding and grasping the laws of economic activities

and business development status of schools, improving the financial management work, and enhancing the financial management level. The importance of studying the optimization measures of financial analysis of colleges and universities is an important part of financial management of colleges and universities. Studying the optimization measures of financial analysis of colleges and universities can provide a more correct and comprehensive understanding of the financial management activities of colleges and universities, grasp the input and use of various resources such as human, material, and financial resources of colleges and universities, analyze the deficiencies in the financial management of colleges and universities, use financial funds more efficiently, and improve the financial management level. In addition, the study of financial analysis optimization measures of colleges and universities can better help financial departments, external investors of colleges and universities, and leaders of colleges and universities understand the financial situation and development trend of colleges and universities and provide factual basis for

scientific [4, 5] decision-making. The current situation and problems of financial analysis in college system include the current situation of financial analysis awareness, the current situation of financial analysis content, the current situation of financial analysis methods, the current situation of financial analysis indexes, etc. With the current situation of financial analysis awareness and existing problems, the leaders do not pay enough attention to it, and the focus of attention is not on the financial analysis work on the one hand, most of the school leaders are education connoisseurs but not financial experts, they know about financial knowledge but not expertise, and they do not know enough about financial analysis and do not pay enough attention to it. On the other hand, the financial leaders of colleges and universities focus on the finance department's work of frequent reimbursement (accounting), tuition collection, budgeting, and declaration of spending and lack the initiative and enthusiasm for financial analysis. In the eyes of most school leaders, financial analysis is to analyze and integrate the data in the financial statements, calculate the relevant ratios and related indicators and, at most, summarize the financial situation with textual explanations. The leaders of colleges and universities do not pay enough attention to it, so it leads to the fact that there are few special financial analysis positions in the financial department of colleges and universities to focus on financial analysis, and it is a common phenomenon that financial staffs write down financial reports part-time every year and make simple and rough analysis and integration according to the report data, which does not achieve the real effect of financial analysis and makes the truthfulness and value of accounting information quality of majesty greatly reduced. The quality of financial analysts is not high on the one hand; the focus of the daily work of the financial personnel in colleges and universities lies in bookkeeping, accounting, and reporting, etc. A lot of energy is put into daily reimbursement work and little financial analysis of the economic activities of colleges and universities, and the focus of the financial personnel themselves is not on financial analysis. On the other hand, financial analysts should not only master down-to-earth professional skills, but also have certain management knowledge and understand statistics [6], probability theory, computer, network, and other aspects of information, and of course, they should be very familiar with the situation of this unit, which puts forward high requirements on the quality of financial personnel. The influence of mutual competition among colleges and universities is fierce. In order to stand out in the competition, schools have to develop continuously, expand the scale of enrollment, and build administrative buildings, teaching buildings, canteens, dormitories, gymnasiums, and other infrastructures, which is not enough only by financial funds and income from educational undertakings. In order to promote the rapid development of the school, attract external investors to invest in the infrastructure construction of the university, and obtain financing in many aspects, the financial analysis of the university system becomes very important. The structure of the index system at all levels of financial risk management assessment of universities is shown in Figure 1.

With the in-depth development of China's university system reform, universities gradually rely on financial analysis in behavioral decision-making, and financial management gradually changes from accounting management to analysis management, and financial analysis plays an increasingly prominent and important role in university management [7–9]. However, in the practical application, there are still many problems in the establishment of indicators and quality of analysis in the financial analysis of universities, which need to be improved continuously. Therefore, it is important to find out the problems in financial management through financial analysis, grasp the law of economic operation, and judge the risk level scientifically, so as to improve the financial management level of universities and promote their scientific development. In this paper, we analyze the design and application of financial analysis indexes of colleges and universities under the new situation, establish a scientific and reasonable financial analysis system, and apply it in order to improve the financial management level of colleges and universities and promote the long-term and rapid development of colleges and universities. Currently, under the influence of planned economic management mode, financial analysis has not attracted enough attention in the financial management of colleges and universities in China, and the management concept of financial analysis is not strong. In financial management, colleges and universities still generally focus on the process of nuclear management and supervision, while lacking postevent analysis and forecast. Even if quarterly or annual analysis is done, it is only an after-the-fact analysis. The financial analysis index system is not sound and perfect, and most colleges and universities do not extend the financial analysis index to the design process, not to mention establishing an efficient analysis model and using special analysis tools, thus failing to provide effective financial information for college behavior decision, not to mention adapting to the needs of rapid development of colleges and universities. At present, there are mainly the following problems in the application of financial indexes in colleges and universities: only simple data listing, lack of systematic analysis and scientific methods; the analysis index system is not comprehensive, and the comparability of relevant data is poor; the design of analysis system is not scientific and lacks accuracy, cannot be quantitative analysis, and is not strong in prediction. Therefore, with the deepening of university reform, China has developed higher education as an industry [10]. Under the new situation, it is especially necessary to carry out scientific design of financial analysis indexes and apply them reasonably. The scientific nature should fully consider the characteristics and nature of the management object and its movement law when designing the financial analysis indexes. The indicators are accurate and complete, neither duplicated nor omitted, independent of each other, and complementary to each other. The purpose of effective financial management is to maximize the economic and social benefits of funds by strengthening management. Financial analysis is for the service of financial management, so the design of financial analysis indexes should be centered on improving efficiency.
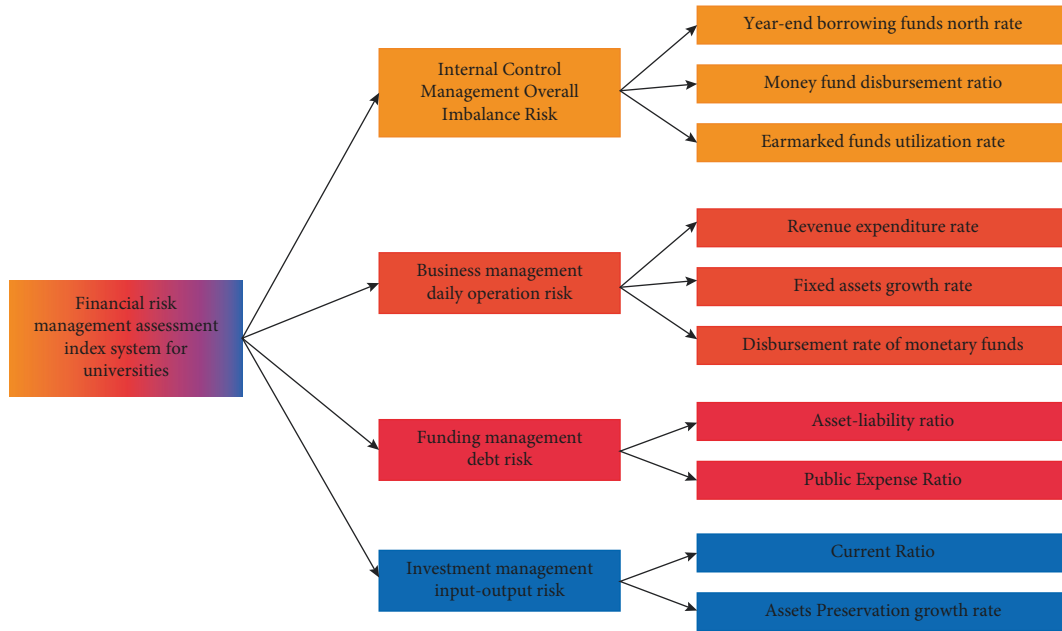
FIGURE 1: Financial accounting work system in colleges and universities.

Practicality the financial analysis index system of universities is a complex system with many contents and complicated relationships. When designing the analysis indexes, we should focus on each of them according to the evaluation objectives and tasks, so that the evaluation results of these indexes have practicality. The financial indicators established by comparability should focus on the interconnection, comprehensive, and developmental view and comparability and avoid one-sidedness, so as to draw correct analysis conclusions. Dynamic accounting and education management are constantly changing with the economic situation, and the design of financial analysis index system should also be updatable. If we do not consider from the dynamic nature, we cannot eliminate the influence of chance factors and lose the meaning of analysis and evaluation [10]. Reasonable determination of the weight of each financial analysis index and its ratio of colleges and universities makes the weight and ratio of each index according to the relevant requirements of the department and combined with the actual situation or according to the needs of their own management and development. The standard ratio of financial indicators of universities in recent years can be calculated by referring to other institutions to determine the specific weighting ratio of our indicators or by using arithmetic average. Reasonably determine the upper and lower limits of each single index score. Relevant authorities or universities should make reference to the weight and standard ratio in the financial analysis index system of universities to set the upper and lower limits according to their own situation when formulating analysis indexes, and universities should not be lower or higher than the upper limit when conducting financial evaluation. Reasonable determination of financial analysis standards of colleges and universities is carried out through reasonable analysis after the standards are formulated, and after the analysis,

excellent, good, qualified, and unqualified are derived, and the results of financial evaluation of colleges and universities can be obtained by substituting them into the relevant score intervals, which are clear and operable.

Dynamic neuron is an optimization method developed in recent years. It adopts the basic idea and structure of dynamic programming and draws on the achievements and methods of artificial neural networks, computer simulation, artificial intelligence, and other fields to optimize the system itself and improve the performance of the system through the simulation of the actual system. It adopts computer simulation and function approximation [11], which can effectively solve the long-standing "dimensional crisis" and has a broad application prospect in manufacturing, communication, logistics, and control engineering fields. Dynamic neuron is a novel artificial neuron model, whose supervised learning aims at stimulating a string of dynamic neurons that express specific information through precise temporal coding, hence the term dynamic neuron learning. The existing dynamic neuron learning methods are reviewed and compared for the significant application value of single-neuron dynamic neuron learning, various theoretical bases, and many influencing factors. The basic concepts of dynamic neuron model and dynamic neuron learning are first introduced; then the typical dynamic neuron learning methods are introduced in detail, and the theoretical basis and synaptic weight adjustment of each method are pointed out; finally, the performance of these learning methods is compared experimentally, and the characteristics of each method are summarized systematically. The results of this study contribute to the comprehensive application of dynamic neuron learning methods [12].

At present, most universities' financial risk identification mainly relies on managers' experience judgment, which is highly subjective and the accuracy varies from person to

person, and it is not easy to discover potential financial risk problems, and it also encounters bottlenecks in promoting the use and publicizing the identification results. Therefore, it is necessary to study a set of financial risk identification models for universities. At present, domestic and foreign scholars have established four main models in financial risk identification: univariate model, multivariate linear discriminant analysis, logistic regression model, and artificial neural network model, which still have shortcomings [13]. The dynamic neuron model has become a hot spot for research because it has no hypothesis requirement for samples and the model has strong fault tolerance, learning ability, and error correction ability, especially in the field of financial risk early warning of listed companies, which has made great progress. However, the research on the quality assessment method of financial accounting work of universities with dynamic neuron model is still in the initial stage. Therefore, this study tries to establish a set of index system for the quality assessment of financial accounting work in colleges and universities from the characteristics of colleges and universities and construct a financial risk identification model for colleges and universities based on dynamic neuron method to achieve the purpose of monitoring and assessment of financial risks in colleges and universities. The main contributions of the model proposed in this paper are as follows:

(1) Starting from the characteristics of colleges, this paper establishes a set of quality evaluation index system of financial accounting in colleges and universities

(2) This paper constructs a university financial risk identification model based on dynamic neuron method, which has a higher recognition rate than traditional machine learning algorithm and static neural network method

(3) The experimental results show that this model can monitor and evaluate the financial risk of colleges and universities, which lays a foundation for relevant research

2. Related Work

2.1. Financial Accounting in Higher Education. Financial analysis is a quantitative reflection of financial data, often only the amount of revenue, expenditure, budget execution rate, and other data, but with the lack of consideration of the impact of changes in revenue, resulting in low budget execution rate, the "quality" of economic operation is not enough to explore the connotation of this emphasis on quantity. This kind of analysis that emphasizes "quantity" but not "quality" inevitably leads to a lack of foresight and guidance in financial work. The financial analysis is not only to conduct microanalysis of our internal finance, but also to conduct horizontal comparison with other institutions of the same level and carry out macroanalysis. The financial analysis of many institutions is based on the description of accounting statement data, the data description of the basic situation of income and expenditure, and the lack of

comparison with the direction, amount, and magnitude of the increase or decrease of the relevant indicators of schools of the same scale of operation or schools with similar operation characteristics, the lack of understanding of the financial situation of the school and the trend of teaching changes, and the lack of reasonable evaluation of their competitive advantages [14]. Neglect of predictive analysis of the results of colleges and universities at the end of economic activities and then regularly has seriously restricted the development of colleges and universities. Colleges and universities should pay more attention to the prior analysis to avoid the drawbacks in the previous work. Otherwise, with the deepening of the reform of college running system, it is difficult for the colleges and universities with unreasonable allocation of educational resources and unsharp professional characteristics to transform into application-oriented universities. Under the new situation, the financial information required for the survival and development of colleges and universities must rely on financial analysis to provide in advance and reduce the competitive pressure of colleges and universities. Neglecting risk factor analysis: Although colleges and universities are nonprofit organizations, they are also in the market economy environment, where the risks and opportunities of running schools coexist and the inputs and benefits go together. Therefore, the financial analysis of colleges and universities only stays on the analysis of some deterministic factors, which has significant hidden dangers. Therefore, colleges and universities should strengthen financial management and strengthen financial risk analysis to prevent major decision-making mistakes from occurring. With lack of cost accounting analysis at present, the economic operation of colleges and universities mainly relies on the income and expenditure standards stipulated in the budget to control the economic activities of each department, so that each department can make full use of educational resources while achieving the set goals. Although the budget control makes the cost expenditure strictly restrained, it ignores the actual situation of various consumptions in economic activities. Therefore, colleges and universities should collect and allocate all kinds of consumptions that occurred in the process of education and teaching, so that the budget and cost can be effectively connected. At the same time, it should also implement the responsibility system for the value preservation and appreciation of state-owned assets and economic cost efficiency indicators and try out the economic responsibility system in a planned way. However, the cost accounting control of colleges and universities is not perfect, the consumption of common resources cannot be measured, teaching work or nonteaching work is costed according to uniform standards, resulting in the lack of authenticity of cost accounting, the educational resources cannot be optimally allocated, and it is difficult to control the phenomenon of occupying and idling resources. Paying attention to the analysis of financial development potential, as a nonprofit institution, the budgeting principles of universities are to balance income and expenditure, keep expenditure within the limits of revenues, and strictly prohibit overspending. The preparation and execution of school budget should be

carried out under this principle, and the red line of "operating in debt" and "deficit budget" should not be touched. Under the situation, it is one of the important indexes in the financial performance assessment of colleges and universities to measure the ability of colleges and universities to bear liabilities and risks and to comprehensively assess and evaluate the financial development potential [14]. The specific contents of the appraisal should include the following aspects: the status of accumulated external liabilities of colleges and universities and the ratio of liabilities, the ratio of annual total income and expenditure of schools, the net balance of deposits at the end of the year, the size of lending and temporary payments and their proportion, the assessment of business risks of school-run industries and the ratio of assets and liabilities, etc. In the new era, the essential task of colleges and universities is to cultivate application-oriented talents, meanwhile, strengthen scientific research and transformation of scientific research achievements, and adapt to the development trend of integration of production, learning, research, and application. Therefore, the performance appraisal of colleges and universities cannot be simply equated with the simple economic benefit appraisal but needs to be combined with various achievement appraisals after quantifying the work of colleges and universities as well as the simple economic benefit appraisal of less investment and more production, replacing the simple benefit appraisal with performance appraisal, and summarizing and analyzing the social and economic benefits with performance indexes. The analysis of university running efficiency is a comprehensive work with the characteristics of science and complexity. When analyzing and evaluating based on the above-mentioned indicators, the staff must integrate relevant statistics, funding arrangements, financial plans, fixed indicators, operations, and other relevant information, carefully sort out, compare, and study and analyze all economic activities and achievements of the university, and follow the principle of comparability and make precise calculations for each analysis indicator of the relevant year. Based on this, we compare and analyze the efficiency indicators between different years to reflect the dynamic changes of the efficiency of the university. Improvement of budget expenditure execution analysis of the expenditure budget of colleges and universities includes the expenditure for carrying out teaching, scientific research, management, and other activities in the year. From a macroperspective, the expenditure budget of colleges and universities can be divided into basic expenditure budget and project expenditure budget. If further subdivided, the basic expenditure budget, for example, can be further divided into the budget for personnel expenses and the budget for daily utility expenses. Among them, the budget of personnel expenses can be further divided into salary and welfare expenses, subsidies to individuals and families, including basic salary, allowance, social security contribution, performance salary, and other salary and welfare expenses; the budget of daily public expenses includes office expenses, printing expenses, consulting expenses, utilities, post and telecommunications, heating expenses, and property management expenses. If the university has implemented the

operation mode of "group management," the above detailed item budget can be further subdivided, and the detailed unit budget can be listed by responsible unit. Analyzing the budget execution rate of each department and preparing the analysis table of expenditure budget execution is an effective way to analyze the execution of university expenditure budget. Taking into account the characteristics and actual situation of colleges and universities, when analyzing the budget execution of each expenditure, it is necessary to analyze and evaluate both the absolute number of overspending or saving of each item and the relative number of the degree of budget execution of each funding item. For the budget execution of each expense item, the university should analyze and evaluate the above two aspects comprehensively to get an objective and fair analysis and evaluation result. At the same time, when analyzing the budget execution of expenditure, we should pay attention to the following contents: the data analysis from the budget execution rate is one-sided, not because the budget execution rate is low, the work of saving and use of funds is good, the budget execution rate is high or the expenditure exceeding the budget is the bad work of saving, which will cause the inaccurate results of budget execution of expenditure, and universities need to evaluate the overspending or saving of each item of funds according to the actual situation of using funds. The financial system of the university is shown in Figure 2. The universities need to evaluate the overspending or saving situation of each expense item according to the actual situation of expense use [15].

Colleges and universities usually use the comparative method and ratio method to analyze and compare the income and expenditure structure of each year, so as to compare the similar indicators with those of institutions with the same scale or similar structure in the same period, compare the differences of indicators, and explore the influencing factors that cause the differences, and make objective and scientific financial analysis. Among them, there are three kinds of comparative analysis methods mainly used by colleges and universities: using expenditure budget as the base, comparing the actual expenditure of funds with the budget, and preparing the corresponding budget expenditure execution table; comparing the current period with the same indicators of the previous period, or comparing the current year with the indicators of the same period of the previous year; level comparison, comparing and measuring the same indicators against similar institutions, and finding out the gaps through comparison and comprehensive thinking. To analyze the trend of financial appropriation income and expenditure of colleges and universities, we use the values of various indicators of financial income and expenditure of a certain period as a fixed base period and compare and analyze the data of financial income and expenditure of other periods or other years with them. Through the analysis of the trend of increase/decrease and magnitude change of the indicators in each period compared with the base period, we can trace the reasons for the fluctuation of the indicators, check the financial accounts, judge the financial situation, determine the economic operation and development trend, and predict the
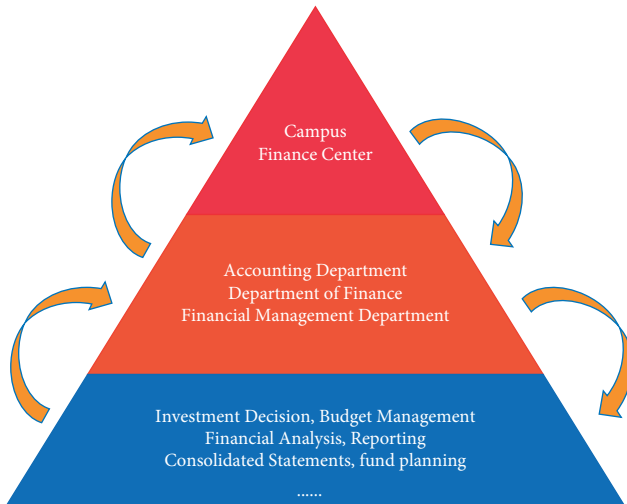
FIGURE 2: University financial system.

possible development results. The indicators with large fluctuations are one of the situations that the trend analysis method focuses on; i.e., it focuses on strengthening the changes of the indicators of special and key parts. By analyzing the trends of the indicators, it is possible to make a general evaluation of the financial income and expenditure and provide a basis for the management to make decisions. The dynamic analysis method takes the quantitative characteristics revealed by the income and expenditure of colleges and universities as the standard, takes the rate of change of economic variables in the process of change in the time series as the basis of analysis, considers the process of change of economic operation from the process of continuous time points, recognizes the law of economic activities, explores the trend of change of economic variables, and judges whether it meets the requirements of strategic development of the university. The fundamental purpose of overall analysis method of financial management is to make limited funds produce greater social and economic benefits. Therefore, when applying the overall analysis method of university finance, it is necessary to follow objective laws, grasp the principle of overall optimization, and pay attention to timeliness, efficiency, and comparability. In this process, colleges and universities should clarify the purpose and take "maximizing efficiency" as the purpose of school economic construction.

2.2. Dynamic Neuron Model. Dynamic neural networks are known as the third generation of neural networks, which operate in a manner closer to real neurons and have stronger performance than traditional neural networks. With pulse excitation time as input and output, dynamic neurons can better simulate biological neurons and thus can be the basis for more effective simulation of artificial intelligence [16]. In the research of artificial neural networks, supervised learning theory has been a central part of the research. Since the physiological basis of various cognitive activities is the sequence of impulses excited by neurons in the human brain, supervised learning of dynamic neurons and networks has

an important application and has become an extremely important part of its research field. Among them, supervised learning with temporal encoding of the excitation of impulses has become a key research direction for researchers to focus on. The supervised learning methods based on temporal encoding of dynamic neural networks can be broadly classified into single-pulse learning and multipulse learning methods according to the number of pulses excited. Obviously, multipulse learning is more in line with the characteristics of biological neuron operation and also has stronger application performance. Supervised learning methods for dynamic neural networks can be classified into single-neuron learning and network learning methods according to the applicable network structure. A general rule for supervised learning methods in terms of network structure and number of pulses is that single-neuron learning methods are generally more capable of learning multiple pulses and can control more pulse excitation moments through learning, while multilayer network learning methods generally control a smaller number of output pulses. A few definitions of the concepts are introduced here first. Supervised learning of dynamic neural networks refers to learning that causes the output neurons to excite a specific sequence of desired output pulses. Time-coded supervised learning of a dynamic neural network means that the output neurons are learned to excite pulses exactly at the specified moment during a period of operation. Time-coded multipulse supervised learning of single dynamic neuron (hereafter referred to as single-neuron learning) refers to a single dynamic neuron with multiple input synapses that is learned to excite multiple pulses at the desired moment, i.e., a sequence of pulses, so it is also referred to as pulse sequence learning in [16]. Dynamic neurons are the basic units that constitute dynamic neural networks, and supervised learning of single dynamic neurons is of great importance as it can better draw on the well-defined synaptic tuning mechanisms of biological neurons, as well as being the basis for implementing supervised learning of more complex neural networks. The process of single-neuron pulse sequence learning is affected by many factors, including the length of the input and output pulse sequences, the excitation rate, and the number of neuronal synapses. The longer the sequence and the higher the excitation rate, the more difficult the learning process; however, many factors also make the pulse sequence learning more flexible with the available information for weight adjustment. Several single-neuron learning methods have been proposed by different researchers. Among them, one class has been proposed based on stochastic models, and such nondeterministic methods are difficult to analyze and compare their learning ability for complex target sequences [17–19]. An impulse sequence learning method is given using a linear algebraic approach, but the theory of this method is too much detached from dynamic neurons and therefore has received less attention from researchers. Also, there has been a review literature on learning methods for dynamic neural networks, but these studies have been presented more broadly and have not specifically focused on impulse sequence learning methods or conducted

experiments to compare the sequence learning ability of different methods. Sequence learning ability is the core performance of pulse sequence learning methods, while methods with stronger learning ability for longer and more complex pulse sequences will have stronger applications and better reflect the characteristics of dynamic neurons.

3. Methods

This section describes in detail the quality evaluation method of university financial accounting based on dynamic neuron model. Firstly, it introduces the specific evaluation indexes of university financial evaluation and then introduces the basic principle of dynamic neural network and how to apply it to university financial evaluation.

3.1. Evaluation Metrics. The index design of financial status analysis of colleges and universities is the performance status of assets, liabilities, and their equity of colleges and universities at a certain point in time. The main indicators are (1) Net asset growth rate = Net asset growth of the current year/Net asset amount of the previous year * 100%. It reflects the efficiency of the university's expenditure and the ability to preserve and increase the value of assets. (2) Ratio of fixed assets formation in current year's expenditure = total capital expenditure in current year/amount of credit incurred in current year's fixed fund * 100%, reflecting the capacity of fixed assets and the ability to preserve the value of fixed assets. (3) Growth rate of fixed assets = Current year's fixed fund growth/Prior year's fixed fund balance * 100%, reflecting the growth of fixed assets. Design of budget execution analysis indicators: Budget execution analysis indicators mainly include the difference between final income and expenditure and budget income and expenditure and income and expenditure structure analysis indicators, etc. (1) Revenue budget completion rate = total final school revenue/total budget revenue * 100%. This indicator reflects the difference between final income and budgeted income, which facilitates the comprehensive and scientific budget management of universities through analysis. (2) Completion rate of expenditure budget = total final expenditure of the university/total budgeted income * 100%. This indicator reflects the difference between final expenditures and budgeted expenditures of colleges and universities. Financial risk situation analysis index designs: The design of financial risk aspect indexes is mainly for risk avoidance. The main ones are as follows: (1) Asset-liability ratio = total liabilities/total assets * 100%. This indicator reflects the ability of universities to repay their debts. The funding source of colleges and universities is mainly financial allocation, and the asset-liability ratio should be controlled at a small ratio. (2) Current ratio = current assets/current liabilities * 100%. This indicator reflects the ability of current assets of colleges and universities to repay current liabilities when they are due. It is more appropriate to keep the ratio at 200%. (3) Cash ratio = Money capital/current liabilities * 100%. This indicator shows the ability of cash to repay current liabilities. (4) Other receivables to current assets ratio = Other

receivables/Current assets ending balance * 100%. Other receivables that have not been collected for more than three years will increase the financial risk of universities. (5) Other payables to current assets ratio = other payables/current assets ending balance * 100%. Take care to avoid financial risk when operating with debt. (6) Loan risk factor = outstanding loan balance (/unrestricted net income * 50% + general fund * 25%): 0.8-1.0 high risk; 0.8-0.6 higher risk; 0.6-0.4 medium risk; 0.4-0.2 low risk; 0.2-0 no risk. Design of indicators for fund utilization analysis: (1) Increase/decrease rate of per capita business expenditure = increase of per capita business expenditure of the current year/per capita business expenditure of the previous year * 100%. This indicator mainly reflects the cost of per capita expenditure of universities. (2) Self-sufficiency rate = (education income + operation income + other income) (/education expenditure + operation expenditure) * 100%. This index reflects the income capacity and the degree of recurrent expenditure of colleges and universities. (3) Growth rate of scientific research funds = the increase of scientific research funds in this year/scientific research funds in the previous year * 100%. This index reflects the ability of universities to serve the society. The evaluation factors of the financial work of colleges and universities are shown in Figure 3.

3.2. Dynamic Neuronal Network. The starting point of artificial neural network research is to mimic the information processing mechanism of biological neural networks. As neurophysiological research continues to advance, it is recognized that biological neural systems have dynamic properties, which are manifested in the delayed nature of synaptic connections, transmission of impulses, and stimulated excitation of cell membranes. In order to simulate dynamic functions such as learning, adaption, memory, and recall, and thus to better reflect the dynamic properties of biological neurons, it is necessary to introduce a feedback link in the mathematical modeling of biological neurons.

Most of the current popular deep networks have the same static reasoning paradigm: once the training is completed, the structure and parameters of the network remain unchanged in the test stage, which limits the representation ability, reasoning efficiency, and interpretability of the model to a certain extent. The dynamic network can adaptively adjust its structure and parameters according to the input samples in the reasoning stage, so it has many good characteristics that the static network cannot enjoy. The dynamic neural unit (DNU) model has both feedforward and feedback weights and then generates output values through the action of nonlinear excitation functions, whose characteristics and structure are different from the traditional static neuron model. The DNU model is shown in Figure 4, where $X_i(k)$ $(i = 1, 2, \cdots, n) \in R^n$ is the input quantity; W_i $(i = 1, 2, \cdots, n) \in R^n$ are the weights; $T \in R^1$ is the threshold; $S(k) \in K^1$ is the input of the linear dynamic system; $a_{ff} = [a_0, a_1, a_2]^T$ is the feedforward weight and $b_{fl} = [b_1, b_2]^T$ is the feedback weight; $V_1(k) \in K^{(-1)}$ is the output of the linear dynamic system; gs is the slope of the
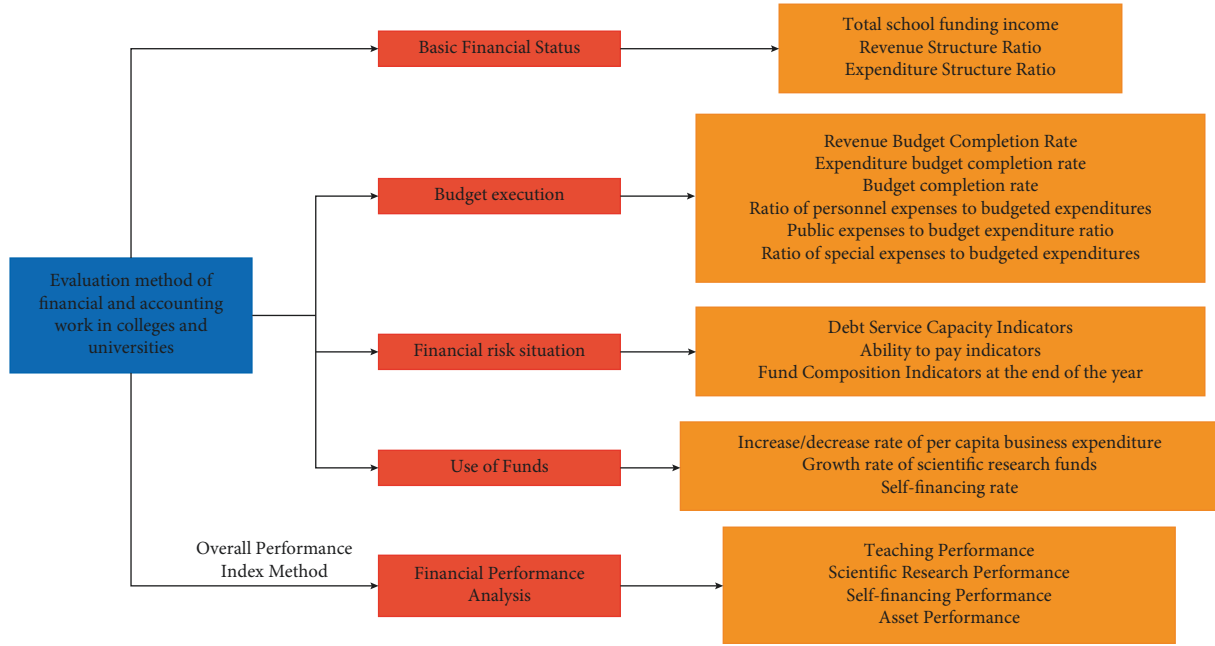
FIGURE 3: Evaluation indicators.

Sigmoid-type nonlinear excitation function curve; $V(k) \in R^1$ is the input variable of the Sigmoid-type nonlinear excitation function; $Y^*(k) \in R^*$ is the output value; k denotes a ...rst moment. The linear dynamic system part of the DNU model is essentially a two-order IIR (Infinite Impulse Response) system, which simulates the reverse circuit in the biological nervous system.

The transfer function of the IIR dynamic system part of the DNU model can be expressed as

$$(k, a\_ff, b\_fb) = \frac{(V\_1(k))}{(S(k)) = [a\_0 + a\_1.\hat{Z}(-1) + a\_2.\hat{Z}(-2)]/[1 + b\_1.Z(-1) + b\_2.\hat{Z}(-2)]}. \tag{1}$$

Equation (1) can also be equivalently expressed in terms of the difference equation as

$$V\_1(k) = -b\_1.V\_1(k-1) - b\_2.V\_1(k-2) + a\_0.S(k) + a\_1.S(k-1) + a\_2.S(k-2) \ (2). \tag{2}$$

From Figure 4, the DNU model treats W_i ($i = 1, 2, \cdots,$ n), T,a_i ($i = 0, 1, 2$), b_i ($i = 1, 2$), g_s as adjustable parameters and modifies their values by learning. The learning algorithm can be optimized by using gradient descent method. Since the DNU model treats the feedback weights as adjustable parameters and modifies their values by learning, this inevitably brings the problem of dynamic stability of the model. In this paper, an improved dynamic neuron model, the CIIRDNU (Constricted Infinite Impulse Response Dynamic Neural Unit) model, is proposed. This model limits the feedback weights b_i ($i = 1, 2$)"," W_i ($i = 1, 2, \cdots,$ n), T, a_i ($i = 1, 2$), g_s as adjustable parameters to modify their values by learning. The learning algorithm is optimized using the gradient descent method. Since the feedback link

variables $b(i = 1, 2)$ are not modified as adjustable parameters in the model, the stability of the model can be ensured with proper initial settings. If $b1 = b2 = 0$, then according to the stability analysis of discrete linear system, $P1 = P2 = 0$, the linear dynamic system part of the model becomes FIR system, which can solve the stability problem of IIRDNU model, that is, another improved dynamic neuron model: FIR-DNU (Finite Impulse Response Dynamic Neural Unit) model. In this paper, traditional pattern recognition methods are combined with neural networks to realize the mapping from sample feature pattern space to pattern category space using feedforward networks of dynamic neuron models. The nonlinear excitation function in the dynamic neuron model is taken as
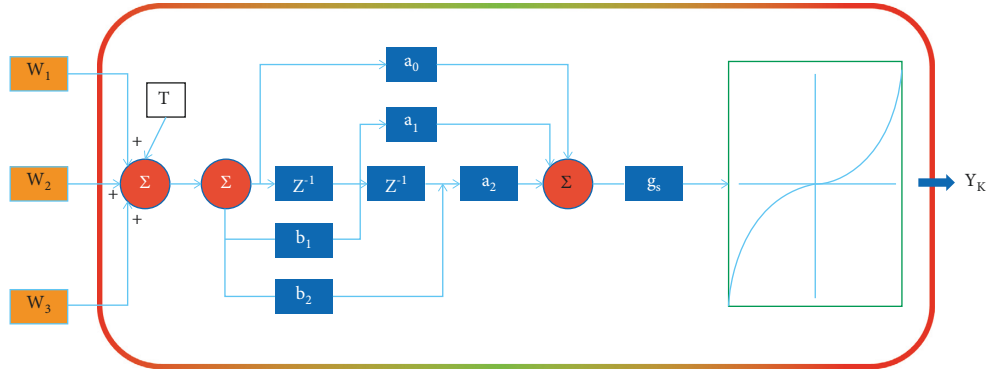
FIGURE 4: Dynamic neuron model.
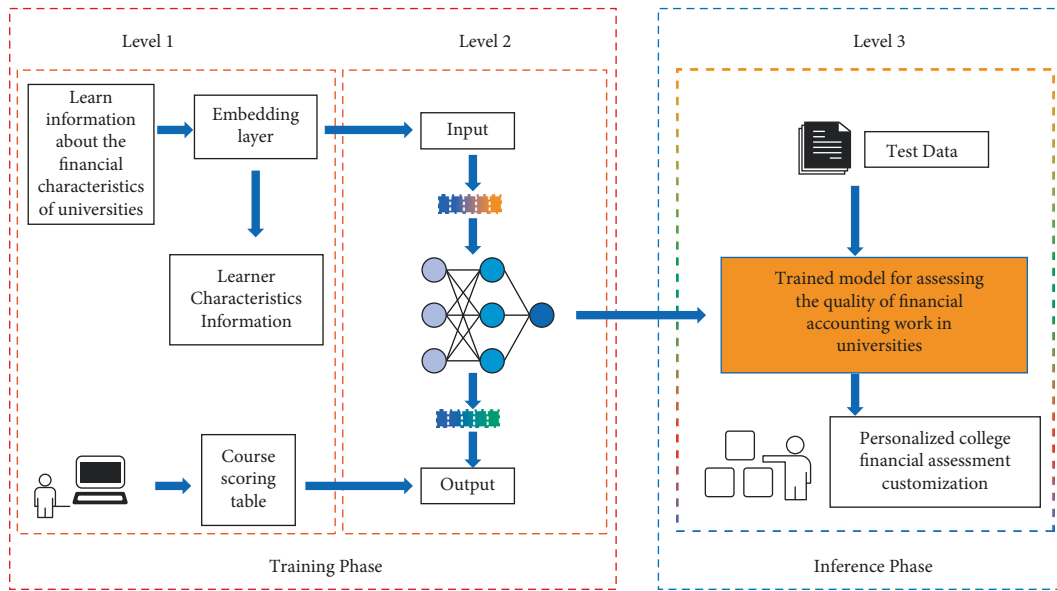
FIGURE 5: Model architecture diagram.

$$Y(k) = \psi[V(k)] = \frac{1}{(1 + \widehat{e} - (V(k)))}. \quad (3)$$

From equation (3), the node output will be equal to 0 or 1 only when $V(k) \longrightarrow \pm\infty$. Therefore, for the binarized target output, the learning goal is considered to be reached as long as the output reaches values close to 0 and 1 at the time of learning. The overall architecture of the proposed model is shown in Figure 5.

4. Results

4.1. Dataset. The financial data of 16 universities (SH1-SH16) in 2018 were collected to construct neural network models, and the BP neural network data mining method was applied to analyze the financial situation of 2 of them (SH15 and SH16) in 2018 for early warning.

4.2. Experimental Setup. The experimental environment configuration information is shown in Table 1.

TABLE 1: Software and hardware configuration information.

Project	Configuration
CPU	Intel Core i7 7700
Memory	128G
GPU	NVIDIA GeForce GTX 2080ti
Operating system	Windows 10
Cuda	Cuda_9.0.176_384.81 with cudnn9.0
Language environment	*Python* 3.6

The number of neurons in the input and output layers of a BP network is determined by the number of dimensions of the input and output vectors. The dimension of the input vector is also the number of influencing factors. Through the analysis of factors affecting the financial status of universities, the article selects three primary indicators such as solvency, operation capacity, and development capacity and 15 secondary indicators such as asset-liability ratio, so the number of neurons in the input layer is 15. In order to refine the financial risk level, the risk level is divided into five levels: giant alarm, heavy alarm, medium alarm, light alarm, and no

TABLE 2: Model output results with actual value.

Colleges and universities	Year	Actual warning value	Network output
SH15	2018	(01000)	(0.1635 1.0000 0.0002 0.0004 0.0064)
SH16	2018	(00010)	(0.0000 0.0003 0.0043 0.9634 0.0051)

TABLE 3: Financial risk identification results.

	Group 1	Group 2	Group 3
Actual output value	1.2457	2.0121	3.2807
Ideal output value	1.0000	2.0000	3.0000
Relative recognition error	25%	1%	9%
Average recognition error		12%	

TABLE 4: Comparison of experimental results.

	Static neural network (%)	SVM (%)	Our model (%)
Recognition accuracy	89	85	98

alarm, so the target output patterns are (00001), (00010), (00100), (01000), and (10000), which are giant alarm, heavy alarm, medium alarm, light alarm, and no alarm, respectively [2, 5]. This gives the number of neurons in the output layer as 5.

Once the network structure is determined, the sample data is used to train through learning rules to improve the network's adaptive capacity. The learning rate determines the amount of change in the weight Figure Neural Network Modeling data stream values in each cycle and is an important factor in the training process. It is generally preferred to choose a smaller learning rate to ensure the stability of learning, and the learning rate is taken as 0.05 in this paper. The article collects the financial data of 16 universities (SH1-SH16) in 2018 as the research sample, and a total of 14 sets of financial indicators of the first 14 universities (SH1-SH14) in 2018 data and financial risk levels as network training samples, and the financial data of the last 2 universities (SH15 and SH16) in 2018 as prediction samples to apply BP neural network for early warning analysis of university financial risks.

4.3. Experimental Results. The results from the network output show that the network achieves the error requirement after 13 training sessions. The network output and the actual risk level of the university are shown in Table 2.

The results of the study show that the BP neural network model has a prediction accuracy of 98%, and the more the important variables for the model, the better the training effect. This shows that the BP neural network successfully forecasts the financial situation of 2 universities (SH15 and SH16) in 2018. The financial risk identification results are shown in Table 3.

4.4. Comparative Experiment. Then, the model in this paper is compared with static neural network and SVM classification algorithm. The experimental method is the same as the previous text. The experimental results are shown in

Table 4. It can be seen that the recognition accuracy of static neural network and SVM algorithm is 89% and 85%, respectively, which are lower than 98% of the accuracy of this model, which verifies the effectiveness of this model.

5. Conclusion

The financial management situation is more a combination of nonfinancial indicators and financial indicators, which can reflect the main measures and effectiveness in terms of budget management, internal control management, asset management, performance management, and talent team building. The description of nonfinancial indicators is an important reference factor to assist financial indicators for trend analysis, which not only is helpful for internal managers to make financial decisions, but also provides important reference information to external supervisors.

Under the new situation, the management system of China's colleges and universities has undergone great changes, and the role of financial analysis in financial management has become increasingly prominent. This paper is practical and valuable for university management by establishing and practically applying the financial index analysis system of universities, and it has certain reference and meaning for other universities to study. However, the financial analysis index system is a dynamic process, which keeps changing with the development of universities. Therefore, the research on the financial analysis index system of universities also needs to be adjusted continuously to meet the development of universities.

In this paper, we will assess the quality of financial accounting work in colleges and universities based on dynamic neuron model. It is studied that feedforward neural network can better reflect the neural computation ability, as long as the appropriate dynamic neuron model is selected (i.e., set the appropriate order n), the single-layer feedforward network established on this basis can solve the problem, and the learning training speed is fast and good. The future research direction is to improve the dynamic

neural network from the perspective of structure and cognitive mechanism, so that it can not only maintain high recognition accuracy, but also adapt to complex and unknown data and have stronger robustness. The dynamic neuron model has the potential for applications that are worthy of further research and development. The assessment results are generally consistent with the actual financial situation of the school, indicating that the financial risk assessment index system has a certain degree of reliability and validity [20].

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] W. E. Bruck and L. Miltenberger, "A school district financial condition assessment system and its application to Pennsylvania school districts," *Journal of Education Finance*, pp. 115–131, 2013.

[2] R. G. Chambers, C. J. Asarta, and E. N. Farley-Ripple, "Gender, parental characteristics, and financial knowledge of high school students: evidence from multicountry data," *Journal of Financial Counseling and Planning*, vol. 30, no. 1, pp. 97–109, 2019.

[3] W. B. Walstad, K. Rebeck, and R. A. MacDonald, "The effects of financial education on the financial knowledge of high school students," *Journal of Consumer Affairs*, vol. 44, no. 2, pp. 336–357, 2010.

[4] S. Zeater, S. I. Benrimoj, F. Fernandez-Llimos, and V. Garcia-Cardenas, "A model for the financial assessment of professional services in community pharmacy: a systematic review," *Journal of the American Pharmacists Association*, vol. 59, no. 1, pp. 108–116, 2019.

[5] P. W. Grimes, K. E. Rogers, and R. C. Smith, "High school economic education and access to financial services," *Journal of Consumer Affairs*, vol. 44, no. 2, pp. 317–335, 2010.

[6] E. O. Arceo-Gómez and F. A. Villagómez, "Financial literacy among Mexican high school teenagers," *International Review of Economics Education*, vol. 24, pp. 1–17, 2017.

[7] C. M. Reich and J. S. Berman, "Do financial literacy classes help? An experimental assessment in a low-income population," *Journal of Social Service Research*, vol. 41, no. 2, pp. 193–203, 2015.

[8] K. Jang, J. Hahn, and H. J. Park, "Comparison of financial literacy between Korean and U.S. high school students," *International Review of Economics Education*, vol. 16, pp. 22–38, 2014.

[9] S. Alvarez, M. Blanquer, and A. Rubio, "Carbon footprint using the compound method based on financial accounts. The case of the school of forestry engineering, technical university of madrid," *Journal of Cleaner Production*, vol. 66, pp. 224–232, 2014.

[10] J. D. Jayaraman and S. Jambunathan, "Financial literacy among high school students: evidence from India," *Citizenship, Social and Economics Education*, vol. 17, no. 3, pp. 168–187, 2018.

[11] S. Shim, B. L. Barber, N. A. Card, J. J. Xiao, and J. Serido, "Financial socialization of first-year college students: the roles of parents, work, and education," *Journal of Youth and Adolescence*, vol. 39, no. 12, pp. 1457–1470, 2010.

[12] D. M. Anderson, K. M. Broton, S. Goldrick-Rab, and R. Kelchen, "Experimental evidence on the impacts of need-based financial aid: longitudinal assessment of the Wisconsin scholars grant," *Journal of Policy Analysis and Management*, vol. 39, no. 3, pp. 720–739, 2020.

[13] C. Erner, M. Goedde-Menke, and M. Oberste, "Financial literacy of high school students: evidence from Germany," *The Journal of Economic Education*, vol. 47, no. 2, pp. 95–105, 2016.

[14] Q. Xu, J. Peng, J. Shen, H. Tang, and G. Pan, "Deep Cov-DenseSNN: a hierarchical event-driven dynamic framework with spiking neurons in noisy environment," *Neural Networks*, vol. 121, pp. 512–519, 2020.

[15] V. Rancic, K. Ballanyi, and S. Gosgnach, "Mapping the dynamic recruitment of spinal neurons during fictive locomotion," *Journal of Neuroscience*, vol. 40, no. 50, pp. 9692–9700, 2020.

[16] V. Fernández-Dueñas, X. Morató, and T. Knöpfel, *Dynamic Recording of Membrane Potential from Hippocampal Neurons by Using a Fluorescence Resonance Energy Transfer-Based Voltage Biosensor[M]//Receptor and Ion Channel Detection in the Brain*, pp. 523–530, Humana, New York, NY, USA, 2021.

[17] A. C. Kaushik, D. Gautam, A. S. Nangraj, D.-Q. Wei, and S. Sahi, "Protection of primary dopaminergic midbrain neurons through impact of small molecules using virtual screening of GPR139 supported by molecular dynamic simulation and systems biology," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 11, no. 2, pp. 247–257, 2019.

[18] H. Yang, H. Wang, and L. Guo, "Dynamic responses of neurons in different states under magnetic field stimulation," *Journal of Computational Neuroscience*, pp. 1–12, 2021.

[19] H. Peng, J. Wang, M. J. Pérez-Jiménez, and A. Riscos-Núñez, "Dynamic threshold neural P systems," *Knowledge-Based Systems*, vol. 163, pp. 875–884, 2019.

[20] C. Liu, X. Jiao, S. Cai, S. He, L. Zhao, and X. Zeng, "Reversible fluorescent probe for visually monitoring the concentration-dependent dynamic correlations among HOCl, H2S, and Ca2+ in neurons," *Sensors and Actuators B: Chemical*, vol. 329, Article ID 129213, 2021.

Research Article

Design of Sports Event Evaluation and Classification Method Based on Deep Neural Network

Shutong Zhao and Jiangang Sun [ID]

School of Physical Education, West Anhui University, Lvan 237012, Anhui, China

Correspondence should be addressed to Jiangang Sun; 12000064@wxc.edu.cn

Large-scale sports events with high-level competition as the main content will have a great impact on the host city whether from the economic level or from the social level. With the improvement of human civilization, people realize that the holding of large-scale sports events not only has a positive impact on the economy and society but also brings some negative effects, such as waste of resources and environmental pollution, which have attracted the attention of the government and investors. Therefore, how to scientifically, comprehensively, and reasonably evaluate large-scale sports events, especially the accurate evaluation of their economic and social effects, has become the focus of attention. The evaluation of large-scale sports events mainly includes two levels: economic and social. Through the specific analysis of the evaluation content and the weight calculation of the evaluation index, the overall optimization of the evaluation of large-scale sports events is realized, and the reference experience is provided for the holding and evaluation of large-scale sports events in the future. Based on this, this article proposes a sports event evaluation and classification method based on the deep neural network. Firstly, on the basis of literature review and field investigation, the evaluation index system of sports events is established. Deep learning models have strong fitting power and robustness and have been applied to many real-world tasks. Then the deep neural network is used to evaluate the holding effect of sports events. The experimental results show that the model has high evaluation accuracy and is of great significance to the supervision and guidance of sports events.

1. Introduction

Sports, as an important part of the social subsystem, has penetrated into every aspect of society. The reemergence of the modern Olympic Movement and the formation of the Olympic economy make large-scale sports events closely related to other social elements such as economy, politics, and culture and play a positive role in the modernization of the city [1]. Since the mid-1980s, the world's major cities have been racing to bid for the large-scale sports event; the main reason is that the city authorities see a big sporting event as an opportunity to exercise and promote the government or the host country or the host city government's administrative ruling ability, thus improving the status and prestige of the host city, expanding urban visibility [2]. For example, Atlanta is known as the world sports center city

because of its advanced sports facilities. Shanghai is also speeding up the pace of building the first-class sports center city in Asia [3].

The holding of large-scale sports events not only improves the political influence of the city but also directly brings great economic effect to the host city and establishes a good self-renewal mechanism and a perfect internal coordination mechanism. The economic driving effect of large-scale sports events on host cities is mainly reflected in stimulating investment, expanding consumption, increasing employment, and optimizing economic structure [4]. Especially in the tertiary industry, tourism, catering, social services, media, and communication industry, the tertiary industry has gradually been in line with the international standards in terms of marketing methods, service technologies, and management concepts. Especially since the

successful holding of the 23rd Los Angeles Olympic Games in 1984, the commercial operation of sports events has brought rich economic returns to the host cities. From the 23rd Olympic Games to the 27th Olympic Games, the direct revenue generated was $227 million, $470 million, $400 million, $100 million, and $450 million, respectively [5]. In addition, competitive sports events can provide more job opportunities for local residents, strengthen the construction of urban infrastructure, and expand economic cooperation and exchanges with various cities.

More and more attention has been paid to the economic and social benefits of sports events by society. The administrative departments of the state and city have been expecting the social and economic benefits of organized sports events while trying to make the society give support to sports events [6]. However, the two sides of things are unavoidable for sports events, and the holding of large-scale sports events will also have a negative impact on the host city. For example, in 1976, Montreal spent a huge amount of money to hold the Olympic Games, which did not bring ideal economic benefits but made it fall into the "Montreal trap." On August 28, 2004, the French newspaper La Tribune revealed that the Greek government had taken on 4 billion euros in debt for the Olympics. For example, during the Sydney Olympic Games, visitors and spectators in the Olympic Village consumed 10,000 cups of beer, 70,000 cups of cola, 3,000 bottles of mineral water, 2,000 pizzas, and 3,000 hot dogs every day [7]. A large number of spectators will leave a lot of waste and garbage for the stadium and the city, which will bring great pressure to the protection of the hosting environment. The combination of declining environmental conditions and a sudden increase in population density has increased the incidence of certain diseases [8].

It can be seen that the impact of large-scale sports events also has positive and negative aspects. Therefore, it is not enough to investigate the economic impetus that sports events bring to the host city. In order to consider the whole process of sports events as the end of a project group, it is necessary to pay attention to the evaluation of society, environment, culture, and other aspects in addition to the evaluation of economic benefits [9]. Therefore, only a comprehensive evaluation of sports events can provide a certain degree of reference for the host region or the host unit and accumulate experience for the holding of subsequent events. The main contributions of this article are as follows: (1) we propose a deep neural network-based evaluation and classification method for sports events; (2) we established a sports event evaluation index system based on the query of related literature and field survey; (3) a deep neural network is used to evaluate the effect of holding sports events; (4) experimental results show that the model has high evaluation accuracy, which is important for the supervision and guidance of sports events.

2. Related Work

The development of project management and evaluation is a result of engineering and project management practice, which mainly originated in the construction industry in the 1930s. With the progress of society and the development of modern technology, the application of related theories of project management and evaluation is also expanding. It can be said that the project management evaluation has gone through a long process from experience to science, which has roughly experienced four stages: subconscious project management evaluation, traditional project management evaluation, project management evaluation dissemination and modernization, and modern project management evaluation [10].

2.1. Status of International Research on Evaluation of Sports Events. At present, all countries in the world are very particular about project evaluation in various fields and the application of various periods; the initial project evaluation in foreign countries is just used in aerospace, defense, and chemical fields; as the subject mutual infiltration and the wide application of computer technology, the project of evaluation also gradually developed into an independent discipline [11]. Some developed countries in the world of international financial organizations in the process of domestic and foreign investment found that traditional project financial analysis methods have been unable to the decision-making of the project in order to guarantee the safety and profitability of its investments, project evaluation also from pure financial analysis to focus on the analysis of the economic development of the economy as a whole system [12].

In the process of project evaluation, some countries in Europe and Asia pay attention to economic benefit evaluation and advocate social evaluation of investment projects, including social equity evaluation, social justice evaluation, evaluation of mutual adaptability between projects and the region where they are located and evaluation of sustainable development [13]. Project social evaluation can be roughly divided into four stages: the first is the project financial evaluation stage. It only existed before the 1950s, characterized by the pursuit of profit maximization. The second is the SCBA evaluation stage. Under the influence of Keynesian economic theory, western economists gradually formed a social cost-benefit analysis method suitable for evaluating public projects [14]. It is from the perspective of the national economy, standing in the national standpoint of the evaluation: the third is the project environmental evaluation stage. Since the 1970s, the theories and methods of environmental assessment of investment projects have been formed and developed gradually. The "material-centered" economic evaluation does not reflect the distribution effect and cannot guarantee the optimal choice of the project [15]. Environmental assessment often ignores the assessment of social and economic impacts. Although it can alleviate the major adverse impacts of some projects, it is insufficient in guiding public participation and reducing negative social impacts. The general steps of sports event evaluation are shown in Figure 1.

As a social project, large-scale sports events have become a bright spot in the field of sports. After the commercial operation of the 1984 Los Angeles Olympic Games, the evaluation of sports events was also elevated to a higher
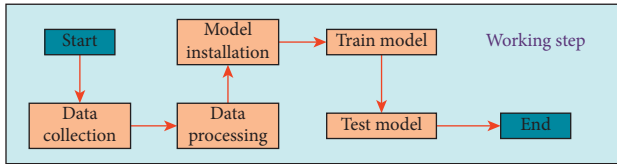
FIGURE 1: The general steps of sports event evaluation.

status [16]. Therefore, the related theories and methods of project evaluation have been widely used in the field of sports, but a professional system has not yet been formed. Judging from the development of the Olympic Games, after the end of the 2001 Sydney Olympic Games, the Sydney Olympic Organizing Committee submitted to the IOC an evaluation report on the overall impact of the Olympic Games on Sydney, which involved about 30 evaluation indicators [17].

2.2. Status of Research on Evaluation of Sports Events in China. At present, in China, project evaluation is widely used in various fields, mainly including public facilities construction project evaluation (highway construction project, urban transportation construction project, etc.), risk management project evaluation, development project evaluation of tourism resources, and project evaluation of ecological environment. The evaluation method is to select the corresponding index in a project and generally adopt a mathematical statistics method to deal with logarithmic excavation or determine the weight of the index, so as to put forward the evaluation system of the project or build the evaluation model of the project [18].

China's sports industry started late until the 1990s to develop. After several decades of development, China's sports industry has already taken shape, and various advanced theories, including project management and evaluation, have also been applied in the sports industry. With the success of Beijing's bid for the 2008 Olympic Games, the evaluation research of large-scale sports events has been pushed to a new height under the background of the Olympic economy [19]. In June 2003, Beijing Organizing Committee for the Olympic Games selected the Renmin University of China as the leader of this project. According to Fang Fu-Qian, a renowned professor at The National People's Congress of China, the overall impact defined by OGGI includes three parts: environmental impact, economic impact, and sociocultural impact [20]. A research team from the Renmin University of China analyzed the 159 indicators suggested by the IOC, and after revising them, the IOC finally confirmed 171 indicators.

The IOC also said individual indicators could be further negotiated. With the start of the 2008 Beijing Olympic Games, the International Olympic Committee requires each Olympic Committee to accurately and comprehensively analyze and study the overall impact of the Olympic Games on the host city or country's social culture, environmental conditions, and economic development and provides about 160 evaluation indicators as a framework [21]. This is the complete evaluation system for the overall effect of large-scale sports events. In the initial 159 indicators provided by the International Olympic Committee, there are 42 items related to environmental evaluation, including the reserve and utilization of water resources, the sewage discharge of Olympic venues, the disposal of household garbage, and the utilization of land. There are 57 indexes involved in economic evaluation, including direct economic benefit and indirect economic benefit. 6.0 items related to social and cultural influence, including public participation consultation, education, social moral construction, and other aspects [22].

2.3. Research Status of Deep Neural Network. Modern convolutional neural network, as one of the most important and active research directions in deep learning, has undergone a long historical development. From the proposal of cybernetics in the 1940s to the development of connectionism in the 1980s, the artificial neural network has been gradually improved. Deep learning was proposed in 2006, which promoted the capability of convolutional neural networks with deep network structure to a new level. For deep learning, one of the most well-known features is that it is inspired by the neural structure of the human brain [23], which is also the historical origin of the term artificial neural network. With this flexible and adaptive structure, mathematicians have built a powerful tool. LeNet, proposed by Lecun in an article published in 1998, is of great significance to modern convolutional neural networks [24], and the basic design of convolutional neural networks in the article is still widely used for reference today. Then in 2006, Geoffrey-Hinton, a professor at the University of Toronto, Canada, published a relevant article in science, marking the rapid development of deep neural network structure [25].

Domestic scholars have also made an attempt to apply deep learning to classification and recognition. Changxin and Gong [26] used the depth confidence network to set the target group layer and scene layer on the basis of the target layer to improve the target detection accuracy through the constraints of the context between the target and the target and the scene context. There is an interaction among the target layer, target group layer, and scene layer to form a supervised network. The network guides the learning of the whole model, adjusts the depth network parameters, and improves the target detection performance. Qi et al. [27] proposed a classification method based on the deep trust networks (DBN) model. Experiments show that compared with support vector machine and traditional neural network methods, the method based on the DBN model can achieve a better classification effect.

3. Establishment of Evaluation and Classification Model

This section describes the sports event evaluation and classification method based on a deep neural network in detail. Firstly, it introduces the determination method of sports event evaluation index weight and then introduces the

basic principle of deep neural network and how to use a deep neural network to evaluate and classify sports events.

3.1. Determine the Weight of Sports Event Evaluation Index.

After the evaluation indicator system is determined, because each indicator has different effects and influences compared with other indicators at the same level, different weights must be determined according to their importance. The weight of the index reflects the relative importance of the whole indicator system. Therefore, the determination of index weight is the most basic evaluation index method. In this study, the analytic hierarchy process was used to determine the weight of each index. According to the analysis of the factors of sports event evaluation, the evaluation criteria of each index are determined, and the evaluation analysis model of a sports event is constructed. The structure analysis model of sports event evaluation is shown in Figure 2.

Pairwise comparison is a method based on comparing two possible combinations of factors to determine the weight value of factors to first-level indicators. According to the principle of analytic hierarchy process (AHP), the importance index of the second round of expert consultation is compared in pairs by using the method of a 9-level grading system, and then all expert opinions are averaged to get a judgment matrix to judge their consistency. Then calculate the relative importance ranking of the index at the same level, namely, the weight of the index in the upper level.

3.2. Basic Principle of Deep Neural Network.

A batch standardization network layer is added to the convolutional neural network after the fusion of the first convolution block and the second convolution block and the feature images extracted from the third, fourth, and fifth convolution blocks [24]. The existence of the BN layer can improve the generalization ability of the convolutional neural network, make the loss function smooth, do not rely heavily on weight initialization, make the network parameters can be randomly initialized, and speed up the network training speed [25]. In the process of neural network training, input values before activation between layers will change in the distribution with the continuous training of the network, and the variation difference of input will accumulate with the deepening of network layers [26], which will cause the gradient disappearing in the process of network deepening. For the gradient descent method used in model training, parameter selection plays an important role in the training process [27]. After the BN layer is added, activation values can be automatically pulled back to normal distribution without deliberately adjusting parameters.

After the third convolution block, the fourth convolution block, and the fifth convolution block pass through the maximum pooling layer, three branches are connected, respectively, and each branch is connected to a global average pooling layer. The features extracted from three convolution blocks are fused together to form fused features for classification, which are connected together to enrich the extracted feature information. The fifth convolution block
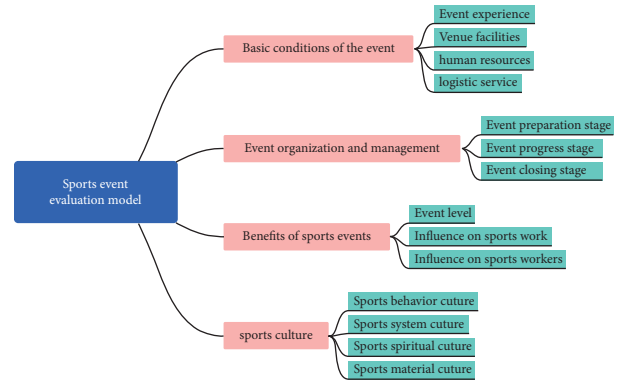
Figure 2: The structure analysis model of sports event evaluation.

extracts the features of the whole band. Finally, the three convolution extracted features are fused together, and the fused features are input into the full connection layer. The fused feature map further highlights each type of feature. The structure of the convolutional neural network is shown in Figure 3.

Average pooling is above the characteristic figure of the extract, similar to the process of convolution, taking the average characteristic figure for dimension reduction results, but the global average pooling layer corresponds to the average of each characteristic in the unit, namely, each characteristic graph output a value, it sets the figure characteristics of the space information and does not require the adjustment of the parameters, reducing fitting. After feature extraction is completed, the extracted features need to be classified and processed, which is the function of the full connection layer, which is located in the last layers of the convolutional neural network. As neurons of the full connection layer need to be fully connected with all neurons of the previous layer, the data volume of the full connection layer is the largest, similar to the BP (backpropagation) algorithm.

When the training set is too small and the feedforward neural network is relatively complex, it is easy for the neural network to overlearn a certain type of data and affect the accuracy, which will lead to overfitting. The basic idea of the Dropout method is to make some neurons in the hidden layer stop working at a certain probability during network training. During the implementation process, the neuron activation function can be reduced to zero at the desired probability. When the Dropout value is set to 0.5, it means that 50 of the 100 neurons in this layer will be set to 0 after the Dropout processing. The neurons randomly set to 0 cannot be the same every time in each iteration update, so the whole process is equivalent to taking the average value of different neural networks. Different random processes produce different overfitting situations so that the overfitting situation is optimized on average. The connection mode of dropout strategy neurons is shown in Figure 4.

In the network, most training parameters are mainly in the full connection layer. By changing the last and second full connection layer of the original neural network to the Dropout layer, the number of parameters can be greatly reduced and overfitting can be reduced. The neural network
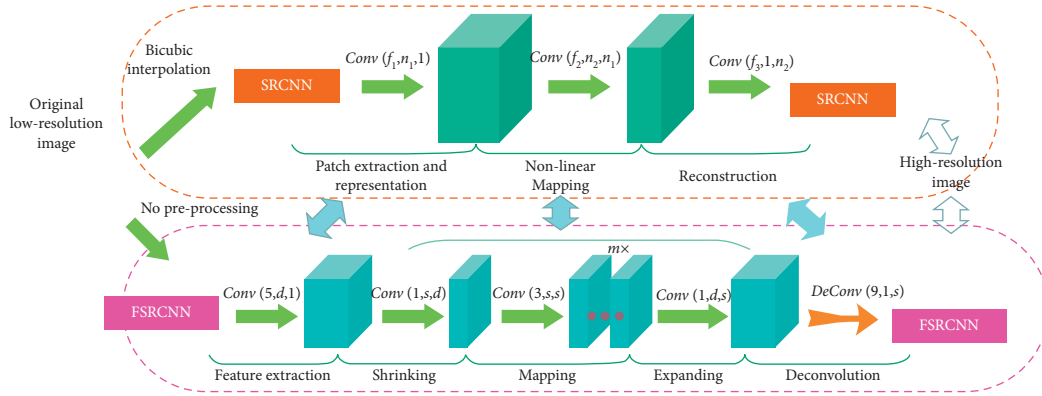
FIGURE 3: The structure of the convolutional neural network.
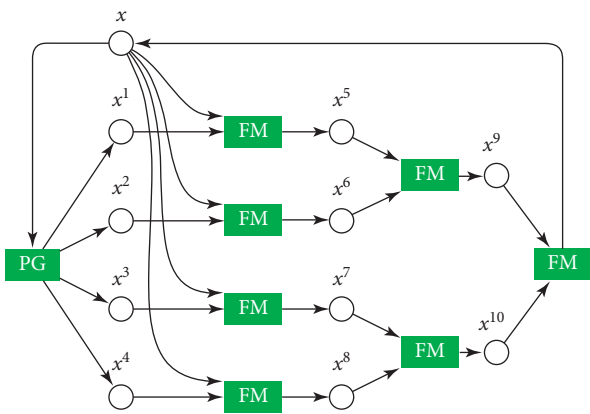
FIGURE 4: The connection mode of dropout strategy neurons.

added different network layers to reduce the original network parameters and reduce overfitting, fused the features of different convolution blocks together to enrich the feature semantics, and optimized the model to improve the accuracy of heartbeat learning for each category of arrhythmias. In the training process of the deep learning classification model, the main purpose is to reduce the error between the actual value and predicted value, that is, to reduce the loss function. The loss function is reduced mainly by adjusting neuron parameters, and the gradient descent algorithm is used to update neuron parameters and reduce the loss function.

A gradient is the derivative of the function at the current position. If a function $f(\theta)$ is convex and θ is an independent variable, then the gradient descent algorithm can be used for optimization, as shown in the following formula:

$$\theta = \theta_0 - \eta \cdot \nabla f(\theta_0), \qquad (1)$$

where θ_0 is the independent variable parameter, η is the learning rate, θ is the updated θ_0, and ∇ is the derivative of this point. The learning rate determines the speed of the network training process and generally takes a small value to avoid missing the global minimum value in the process of gradient update. The learning rate will constantly adjust its own value in actual network training so that training parameters can be updated more quickly.

Backpropagation, including forward process and backward process, after the forward propagation from the input layer to the output layer, and using the prediction results and the real worth to loss function values, to spread after using the loss function values obtained from the front, from the output layer to input layer, network parameter gradient, twice before and after the process, is complete parameter iteration process at a time. The basic principle comes from the basic chain rule of calculus:

$$\frac{dz}{ds} = \frac{\partial z}{\partial x}\frac{dx}{ds} + \frac{\partial z}{\partial y}\frac{dy}{ds}. \qquad (2)$$

The constant updating of model parameters is realized by reducing the loss function continuously through these two algorithms. Generally, n samples are selected as a batch sample during the training of the model, and the error between the predicted value and the real value is obtained through forward propagation operation, and then the parameter is updated by the gradient descent method. The classification of this model is shown in Figure 5.

All samples of different batches were trained according to the principle of no return, and all samples were studied to become an epoch. Gradient propagates backward, updating parameters layer by layer from back to front, which is a batch process. The network output can be fully utilized and the accuracy of network identification can be improved.

3.3. Gradient Optimization of Neural Network. As mentioned in the convolution neural network, a simple introduction of stochastic gradient descent has been proposed and a widely used optimization algorithm has made great progress; in this article's method study phase, the contrast experiment and the analysis of several kinds of the optimization algorithm are conducted. One of the biggest problems of gradient optimization in feature space in the early stage of deep learning is that it falls into local minimum value during the optimization process, and the problem of gradient noise in SGD training also largely reflects the training speed and effect. Hessian matrix and random gradient variance are difficult to control. The formula is as follows:

Euclidean Metric   Mahalanobis Metric
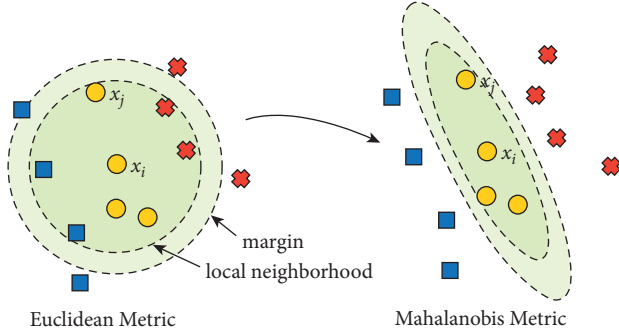
FIGURE 5: The classification of the model.

$$v = \alpha v - \varepsilon \Delta \theta \left(\frac{1}{m} \sum_{i=1}^{m} L\left(f\left(x^{(i)}; \theta \right), y^{(i)} \right) \right),$$

$$\theta = \theta + v,$$

(3)

where v is the velocity term, α is the momentum hyperparameter controlling the gradient decay rate, and its value is between [0,1]. Obviously, the momentum algorithm adds the idea of energy term to the gradient optimization algorithm by introducing the velocity term for gradient update. In the algorithm, the degree can be understood as the exponential decay average value of the negative gradient. According to classical mechanics, the negative gradient is the force acting on a particle moving at a certain speed. The point is assumed to be a unit mass in the feature space so this energy equation can be understood as the momentum of gradient optimization in the feature space. On this point, this article then makes a comparative experiment on such gradient optimization algorithms combined with data. The specific algorithms are as follows:

$$g = \frac{1}{m} \Delta \theta \sum_{i} L\left(f\left(x^{(i)}; \theta \right), y^{(i)} \right),$$

(4)

where g is the gradient, m is the dimension of i-th, L is the function to be optimized, and $y^{(i)}$ is a batch data containing m new sample input. Numerical stability of the algorithm used for decimal division: the final parameter update quantity is obtained through the above equation. The convolution kernels of the mini4 convolutional neural network designed are all $3 * 3$ in size. Small convolution kernels are the development trend of the current convolutional neural network, which can reduce network parameters and reduce computation on the premise of maintaining accuracy. Firstly, there are two convolution layers with 32 convolution kernels. After extracting features from the two convolution layers, the maximum pooling layer is used to reduce dimension and compress features. It is followed by a convolution layer with 64 convolution kernels to continue the feature abstraction through the maximum pooling layer, and finally add the full connection layer. Features extracted from the full connection layer are classified by the softmax function. The gradient optimization diagram of the neural network is shown in Figure 6.

The function of the pooling layer is to further subsample the feature values extracted from the convolutional layer. The subsampling processing of the feature graph can further reduce the data scale, and it has local linear transformation invariance, which improves the generalization processing ability of the convolutional neural network. It can be used to compress images and reduce overwork. There are generally two methods of pooling layer operation: one is the maximum pooling function that takes the maximum value and the other is the average pooling function that takes the average value. After the feature graph is obtained by the convolution operation between data and the convolution kernel, if the linearity is transmitted to the next layer, the linear relationship will result in that no matter how many hidden layers there are in a neural network, the output is the input, and the hidden layer has no effect.

4. Experiments and Results

4.1. Model Training Experiment. Our dataset comes from the organizers of major domestic sports events in China, with a total of 100,000 pieces of data, 7750 pieces for both the test and validation sets, and the remainder for the training set. In the training process of the convolutional neural network, in order to maintain comparability, the same learning rate is set to 0.001, 32 samples are used as a batch, the impulse parameter is set to 0.9, and the dropout value is set to 0.5. The number of iterations is 20. The optimizer is selected as the batch random gradient descent method, and the cross-entropy loss function is selected as the loss function. The experimental environment of this article is as follows: the hardware environment is a Linux system, NVIDIA GTX 2080Ti; the software environment is Python3.5, sklearn0.20.3, and other toolkits. The model loss function diagram and accuracy diagram are shown in Figure 7.

After the model parameter training is completed, the model is saved and the model classification results are verified by the test set. The confusion matrix of convolutional neural network results is shown in Table 1.

As can be seen from the convolutional neural networks, the 4-layer mini4 convolutional neural network is shown in Table 1, and the classification effect is poor. Class L and R heartbeats are identified as more than class, so it is not discussed anymore. Network and improved network VGG-Connect are used as the evaluation objects in the following article. According to the above classification of confusion matrix, the classification model was evaluated using two indexes: sensitivity (Se) and positive predictivity (Pp). The index formula is as follows.

$$Se = \frac{TP}{TP + FN} \times 100\%,$$

$$Pp = \frac{TP}{TP + FP} \times 100\%.$$

(5)

The key to the construction of the evaluation index system of sports education lies in its feasibility in practical application. Through data recovery and statistical analysis,

For factor 2

Stride = 2

The convolution filters can be shared for different upscaling factors

Conv (5,d,1) Conv (1,s,d) Conv (3,s,s) Conv (1,d,s)

DeConv (9,1,s)
Deconvolution layer

For factor 3

Stride = 3
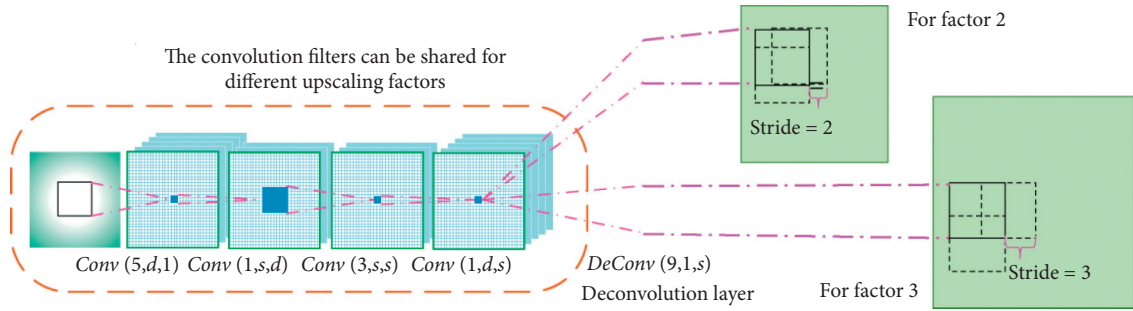
FIGURE 6: The gradient optimization diagram of neural network.

— Loss
— Acc
— Val–Loss
— Val–ACC
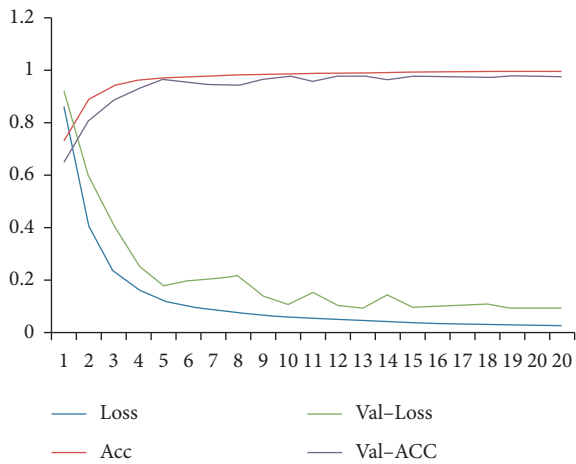
FIGURE 7: The model loss function diagram and accuracy diagram.

TABLE 1: Convolutional neural network confusion matrix.

	A	l	n	r	v	Total
A	585	0	39	2	4	630
L	0	1067	1	0	2	1070
N	24	0	3472	2	2	3500
R	3	0	4	1063	0	1070
V	37	9	56	8	1370	1480
Total	649	1076	3572	1075	1378	7750

the measured results are given, and the feasibility of the index system is preliminarily verified, which provides a reference for the evaluation of sports events in primary and secondary school. Indicators can reflect the situation between the real value and the predicted value of the classification results of the model, and they are used to evaluate the meanings of each part in the classification formula of the model, as shown in Table 2.

According to the identification of the test set, after statistics and calculation, each corresponding index is obtained and drawn as shown in Table 3. Combined with the previous research, analysis, and comparative experiments, the 22-layer deep convolution neural network with inception module is selected for training at this stage. The network performance is evaluated by ten groups of data of each type. Through the error matrix, it can be seen that the network has recognition errors in three categories, but overall, the

recognition accuracy is high, which can effectively identify a single data block in the data.

It can be seen from Table 3 that in the improved VGG convolutional neural network t, the average sensitivity of the classification results of five kinds of abnormalities is 96.74%, and the average positive detection rate is 96.96%. Sensitivity se represents the recall rate of abnormal classification. The higher se indicates the lower probability of missed diagnosis. The positive detection rate PP represents the accuracy rate of abnormal classification. The higher positive detection rate PP indicates the higher correct detection ability. The improved network VGG Connect improves the correct detection ability (average positive detection rate) by 0.35%.

4.2. Suggestions on Improvement Strategies of Sports Events. In the previous section, the evaluation score of the sports event mode is calculated. This chapter will analyze the evaluation results and put forward improvement suggestions. In 2017, more than 2000 sports events were broadcast live. Therefore, the specific analysis of the evaluation results of sports event mode is representative of the sports event live broadcasting platform according to the evaluation of sports business mode and puts forward improvement strategy suggestions for sports event mode, which is of practical significance for the improvement of Internet sports business mode. According to the evaluation results of sports event mode, we can find out the optimization order of sports event mode, that is, find out the places where the sports event mode needs to be improved most in order to improve the value of sports event mode in the shortest time. The calculation formula is as follows: difference value = weight $*$ (1-measure value), where measure value = index score/100. The difference value indicates the impact of the index on the sports event mode. The greater the difference value, the greater the impact of the index on the sports event mode. Therefore, the optimization order of the index can be sorted by the difference value. The calculation results of the difference value of the primary indicators of sports event mode evaluation are shown in Table 4.

By analyzing the characteristics of the actual sample data, it can be found that the classification effect is very stable in the complete data, the recognition accuracy decreases greatly, and the other two types have a large degree of dispersion in the error matrix. However, through in-depth learning of the commonly used top accuracy matrix

TABLE 2: Formula of meaning.

Names	Meaning
TP	The number of numbers that are predicted to be true and actually true
FP	The number of numbers that the model predicts to be true and are actually false
FN	The model predicts false, and the actual number is true

TABLE 3: Neural network partitioning results indicators.

	TP	FP	FN	Se(%)	Pp(%)
A	597	91	33	94.76	86.77
L	1065	7	5	99.53	99.35
N	3430	50	3472	98.00	98.56
R	1061	8	4	99.16	99.25
V	1428	13	56	96.49	99.10
Total	7581	169	169	97.59	96.61

TABLE 4: The calculation results of the primary indicators of sports event mode evaluation.

Primary index	Weight	Measure value	Differences in values	To improve the order
User	0.375	0.830	0.064	1
Service	0.281	0.802	0.056	2
Infrastructure	0.141	0.902	0.014	4
Financial	0.202	0.783	0.044	3

TABLE 5: Error matrix with large error.

Primary index	Weight	Measure value	Differences in values	To improve the order
Relationship	0.226	0.844	0.035	2
User segment	0.079	0.812	0.015	5
Channel	0.070	0.804	0.014	6
Value	0.281	0.802	0.056	1
Cooperation	0.033	0.914	0.003	4
Cost	0.072	0.896	0.006	3

(see Table 5), the classification accuracy still has traces to follow.

The top accuracy error matrix is formed by arranging the soft classification membership degrees of a group of data in descending order. It can be found that although there is a recognition deviation in the item with the highest membership degree, the network can still recognize the correct classification in the two items with the highest membership degree, even in the case of wrong judgment. At this stage, the recognition performance of a deep learning correlation network for data classification is difficult to achieve a completely reliable classification accuracy. Therefore, in the actual accuracy evaluation process, smooth interpolation is made for the two soft classification items with the highest membership as a joint reference to modify the final recognition method.

5. Conclusion

Based on the principles of sports event evaluation and school sports theory, combined with the actual situation of China's current sports event development, through field investigation and interview, this article puts forward the theoretical framework of establishing sports event evaluation index system. Focusing on the goal of sports event evaluation, this article analyzes the situation of

sports events from four aspects: basic conditions of sports events, organization, and management of sports events, benefits of sports events, and sports culture. Based on new technologies such as convolution neural network with deep structure, the construction of an automatic evaluation technology system has great theoretical and practical significance. Based on this, this article proposes a sports event evaluation and classification method based on deep neural network. Starting from the convolution neural network, combined with the existing technical basis, this article analyzes the applicability of this method through comparative experiments and makes targeted fine-tuning and construction from the network model. Combined with the data characteristics, the dataset suitable for network training and the preprocessing strategy suitable for network evaluation are constructed. The convolution neural network technology is effectively combined with the event evaluation method. The experimental results show that the proposed model has high classification accuracy.

This model can be used for the horizontal comparison of sports events in different parts of the country, looking for their shortcomings, and understanding the development of sports events in different parts of the country. It can also be used to compare various indicators and find

problems so as to better improve the infrastructure construction of sports events, standardize the organization and management of sports events, and improve the benefits of sports events.

However, this method also shows some deficiencies. The following combs and prospects some deficiencies of this article. Many methods involved in this paper are realized automatically by the program. With the deepening of research in the future and combined with the existing research foundation, it is expected to form a relatively complete set of automatic detection software. In the future, we plan to develop cross-modal learning applications based on natural language processing and deep vision for the evaluation and classification of large-scale sports events.

Data Availability

Some or all data, models, or codes generated or used during the study are available from the corresponding author by request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. Koenig-Lewis, Y. Asaad, and A. Palmer, "Sports events and interaction among spectators: examining antecedents of spectators' value creation," *European Sport Management Quarterly*, vol. 18, no. 2, pp. 193–215, 2018.

[2] B. Masanovic, "Attitudes of consumers from Autonomous Province of Vojvodina toward advertising through sport in relation with the frequency of watching sports events," *Sport Mont*, vol. 16, no. 3, pp. 91–96, 2018.

[3] D. Stupar, J. Gardasevic, J. Gardasevic, and B. Masanovic, "Attitudes of consumers from Educons university toward advertising through sport among the frequency of watching sports events," *Journal of Anthropology of Sport and Physical Education*, vol. 2, no. 3, pp. 131–136, 2018.

[4] D. A. Harjito, M. Alam, and R. A. K. Dewi, "Impacts of international sports events on the stock market: evidence from the announcement of the 18th asian games and 30th southeast asian games," *International Journal of Sport Finance*, vol. 16, no. 3, pp. 139–147, 2021.

[5] C.-V. Priporas, C. A. Vassiliadis, N. Stylos, and A. K. Fotiadis, "The effect of sport tourists' travel style, destination and event choices, and motivation on their involvement in small-scale sports events," *Event Management*, vol. 22, no. 5, pp. 745–765, 2018.

[6] S. Moon, J. Lee, D. Nam, W. Yoo, and W. Kim, "A comparative study on preprocessing methods for object tracking in sports events," in *Proceedings of the 2018 Twentyth International Conference on Advanced Communication Technology (ICACT)*, pp. 460–462, IEEE, Chuncheon, Republic of Korea, February 2018.

[7] L. A. Rapoport, E. V. Kharitonova, and A. S. Markova, "Major sports events: regional development catalyzing effects," *Theory and Practice of Physical Culture*, no. 7, pp. 67–69, 2021.

[8] P. Schlemmer, M. Stickdorn, E. Kristiansen, and M. Schnitzer, "A mixed methods stakeholder satisfaction study of sports events based on the case of the 2016 international children's games," *Journal of Convention & Event Tourism*, Routledge, vol. 23, pp. 1–18, 2021.

[9] S. Shultana, M. S. Moharram, and N. Neehal, "Olympic sports events classification using convolutional neural networks," *Proceedings of International Joint Conference on Computational Intelligence*, Springer, Singapore, pp. 507–518, 2020.

[10] L. Strzelecki and T. Czuba, "Sports events as an effective way of brand marketing communication using the example of the Orlik sports facilities," *Baltic Journal of Health and Physical Activity*, vol. 10, no. 3, pp. 124–135, 2018.

[11] Z. Hao, X. Zhang, V. P. Singh, and F. Hao, "Joint modeling of precipitation and temperature under influences of El Niño Southern Oscillation for compound event evaluation and prediction," *Atmospheric Research*, vol. 245, Article ID 105090, 2020.

[12] A. Scheu and H. Preuss, "Residents' perceptions of mega sport event legacies and impacts," *German journal of exercise and sport research*, vol. 48, no. 3, pp. 376–386, 2018.

[13] A. Thomson, G. Cuskelly, K. Toohey, M. Kennelly, P. Burton, and L. Fredline, "Sport event legacy: a systematic quantitative review of literature," *Sport Management Review*, vol. 22, no. 3, pp. 295–321, 2019.

[14] J. M. Vegara-Ferri, D. Ibáñez-Ortega, M. Carboneros, J. M. López-Gullón, and S. Angosto, "Evaluation of the tourist perception of the spectator in an eSport event," *Publicaciones*, vol. 50, no. 1, pp. 371–384, 2020.

[15] G. T. Trail and B. P. McCullough, "Marketing sustainability through sport: testing the sport sustainability campaign evaluation model," *European Sport Management Quarterly*, vol. 20, no. 2, pp. 109–129, 2020.

[16] J. A. Rice, M. E. Hambrick, and T. C. Greenwell, "Mass participant sport event brand associations: an analysis of two event categories," *International Journal of Sport Management and Marketing*, vol. 19, no. 5-6, pp. 330–351, 2019.

[17] Q. Zeng, J. Liu, D. Yang et al., "Machine learning based automatic sport event detection and counting," in *Proceedings of the 2021 Seventh IEEE International Conference on Network Intelligence and Digital Content (IC-NIDC)*, pp. 16–20, IEEE, Beijing, China, November 2021.

[18] T. Suehiro, T. Kobayashi, O. Takyu, and Y. Fuwa, "Performance evaluation of efficient collection method of sensor information in wireless sensor networks for event detection," in *Proceedings of the 2021 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 371–376, IEEE, Jeju Island, Republic of Korea, October 2021.

[19] A. E. Bauman, M. Kamada, R. S. Reis et al., "An evidence-based assessment of the impact of the Olympic Games on population levels of physical activity," *The Lancet*, vol. 398, no. 10298, pp. 456–464, 2021.

[20] J. Zhang, "Researching the LPP for the 2008 Beijing Olympics," *Language Policy and Planning for the Modern Olympic Games*, pp. 40–61, De Gruyter Mouton, Berlin, Germany, 2021.

[21] K. Lai, "Influence of event image on destination image: the case of the 2008 Beijing Olympic Games," *Journal of Destination Marketing & Management*, vol. 7, pp. 153–163, 2018.

[22] P. Yang, X. Sheng, Y. Zhao, and L. Zhu, "Evaluation of ecological civilization development in the post-Olympic times," *Applied Ecology and Environmental Research*, vol. 17, pp. 8513–8525, 2019.

[23] G. E. Hinton and T. Shallice, "Lesioning an attractor network: investigations of acquired dyslexia," *Psychological Review*, vol. 98, no. 1, pp. 74–95, 1991.

[24] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[26] G. Chang-Xin and S. gong, "High-resolution remote-sensing image object detection based on deep learning," *Surveying and Mapping Bulletin*, no. S0, pp. 108–111, 2014.

[27] L. Qi, Y. Dou, X. Niu, and J. Xu, "Remote-sensing image classification based on the DBN model," *Computer Research and Development*, vol. 51, no. 9, pp. 1911–1918, 2014.

Research Article

Simulation of Logistics Delay in Bayesian Network Control Based on Genetic EM Algorithm

Pengliang Qiao [1,2]

[1]*School of Management, Guangzhou College of Technology and Business, Foshan 528138, Guangdong, China*
[2]*School of Economics and Management, Beibu Gulf University, Qinzhou 535011, Guangxi, China*

Correspondence should be addressed to Pengliang Qiao; qpliang@gzgs.edu.cn

With the continuous development of e-commerce, the logistics industry is thriving, and logistics delays have become an issue that deserves more and more attention. Genetic EM algorithm is a genetic EM algorithm that is an iterative optimization strategy algorithm that can be used to solve the high-quality algorithm of travel problems with many nodes. Bayesian network (BN) is a network model based on probabilistic uncertainty. This article aims to study the probability of many factors that cause logistics delays to construct an algorithm model to control or reduce logistics delays. This paper constructs an EY model (That is the abbreviation of BN model based on genetic EM algorithm) based on the genetic EM algorithm, and conducts related simulation experiments based on the model to verify the accuracy and feasibility of the model. The experimental results of this paper show that the calculation efficiency of the EY model is significantly improved, and the actuarial accuracy is as high as 98%, which can effectively control logistics delays.

1. Introduction

1.1. Background and Significance. The concept of "logistics" varies from country to country and time to time. Generally speaking, logistics means the circulation of goods. The earliest logistic activity was recorded in the UK. The purpose at that time was to deliver the products to wholesalers, retailers and users nationwide in a timely manner. During World War II, the United States took military needs as the starting point and first adopted the term "logistics management" to refer to the supply of weapons during the war. In today's e-commerce era, the global logistics industry is experiencing new developments. The central goal of modern logistics services is to meet customer needs with the lowest comprehensive cost in the entire logistics process. However, with the rapid development of e-commerce, online shopping users' complaints about logistics delays have increased year by year. After analyzing the reasons in detail, it is found that the current solution to logistics delays is an urgent and resource-intensive emergency solution. This plan is incompatible with the normalization and globalization of e-commerce logistics packages. Therefore, it is necessary to have a deep understanding of e-commerce logistics, and formulate accurate measures based on scientific law analysis.

With the popularization of e-commerce now, logistics not only brings a lot of inconvenience to transportation, but once it is delayed, it will be subject to complaints from many customers and even large claims from partners. Studying how to control logistics delays and effectively avoid losses caused by logistics can not only reduce the complaint rate and improve customer satisfaction but also gain the trust of partners and win more cooperation opportunities, which is of great commercial significance.

1.2. Related Work. With the continuous development of the Internet of Things technology and the continuous prosperity of the logistics industry, more and more adults have conducted relevant research and exploration on logistics delays. With the development of the logistics industry, road traffic has become more and more congested. Due to the continuous changes in road conditions, the unstructured traffic

chaos, the lack of lane discipline, and the wide variety of vehicles, there is a special need to develop a new type of traffic monitoring system. Chaudhary S uses dynamic BN to learn the pattern of road conditions and predict future road traffic conditions within a specific time delay. The predicted road condition information can be used for traffic planning. The method proposed by Chaudhary et al. is computationally lightweight, but very powerful and efficient. The algorithm has also been tested against various weather conditions. Chaudhary et al. uses Synchro Studio simulator (a traffic simulation software) to verify the algorithm, the average accuracy rate is 95.7%, for real-time video, the experimental accuracy rate is 84% [1]. Subjective detection is a natural language processing task designed to remove "facts" or "neutral" content from online product reviews, or objective text that does not contain opinions. Chaturvedi extends the extreme learning machine (ELM is a type of machine learning system or method based on feedforward neural network, suitable for supervised learning and unsupervised learning problems) paradigm to a new framework that uses the features of Bayesian and fuzzy recurrent neural networks to perform subjective detection. In particular, BN is used in the traditional ELM configuration to establish a network of connections between hidden neurons. The experimental results of Chaturvedi et al. verified the ability of the proposed framework to deal with standard subjective detection problems, and proved its ability to solve the portability between languages in translation tasks [2]. The multidimensional BN classifier is a topology-constrained BN, specifically used to classify data instances into multidimensional. Bolt and van der Gaag introduces the concept of balanced sensitivity function, in which multiple parameters are related in function by the odds ratio of the original value and the new value. These functions provide appropriate inspiration for adjusting the multidimensional classifier (It is a classification model that describes the dependency between class variables and class variables, class variables and attribute variables, and attribute variables and attribute variables) and output probability. The practical application of Bolt's classifier in the veterinary field proves the practicality of the heuristic method [3]. In ship reliability engineering, reliability evaluation is an important step to determine whether a ship has an acceptable level of reliability. Current reliability assessment methods are based on static logic. However, warships are typical power systems, and reliability cannot be evaluated by the above methods. Liang et al. proposed a new reliability evaluation method based on dynamic BN and numerical simulation. It overcomes the limitations of analysis methods and multi-level integration methods, and provides an effective means for evaluating ship reliability [4]. Based on dynamic BN with incomplete time slices and mixed Gaussian output, Zhang et al. proposed a data-driven method for predicting process failures with unknown models. According to the requirements of failure prediction, this method constructs an incomplete time slice BN of unknown future observation node. The method proposed by Zhang et al. is suitable for continuous stirred tank reactor system. The results show that even if some measured values are lost, the method proposed

by Zhang et al. can effectively predict and identify faults and estimate the remaining service life of the process [5]. The original BN model did not have the expected variable hierarchy. When modeling the relationship between product sales in the retail industry, it is best to consider the hierarchy of items. Kawabe et al. proposed a hierarchical BN model based on the maximum and minimum hill climbing learning algorithm and a new learning method. The method proposed by Kawabe et al. focuses on the product hierarchy and allows the construction of a lower layer that considers the causality of the upper layer. In addition, Kawabe et al. investigated case studies using actual hierarchical data of consumer purchases and demonstrated the proposed model using simulation analysis [6]. Legal scholars have put forward examples of reasoning problems such as the twin problem and the poison paradox (that is, both the idea of drinking medicine and the cancellation of the idea of drinking medicine must be generated, which violates the law of contradiction) to prove the limitations of probability theory in legal reasoning. Specifically, such questions are intended to show that the use of probability theory may lead to legal paradoxes. Therefore, these problems have seriously affected the application of probability theory in law, especially Bayes' theorem. If various related hypotheses and evidence are correctly described in BN, the paradox disappears. Jacob demonstrated this with examples of food trays, the abuse paradox, and the murder problem in small towns. The resulting BN provides a powerful framework for legal reasoning [7]. However, the above research either lacks actual testing, or the calculation process is too complicated or even accurate, and it is currently difficult to pass.

1.3. Innovation. (1) The use of genetic EM algorithm can perform convergence calculations (it refers to the solution that has been numerically calculated, and if it continues to iterate, the result will not change) on unknown parameters, which can reduce logistics delay time. (2) Using BN to represent the graphical model of the connection probability of a set of variables, it provides a method for calculating logistics delays. (3) The EY model combined with the genetic EM algorithm and BN can greatly reduce the calculation error. (4) Comparing the EY model in this article with the traditional calculation model can better reflect the advantages of this model.

2. BN Model Based on Genetic EM Algorithm

2.1. Genetic EM Algorithm

(1) The concept of genetic EM algorithm. The full name of genetic EM algorithm is Electromagnetic-like Algorithm, which is similar to electromagnetic mechanism algorithm. This algorithm is a new global improvement algorithm based on the meta-heuristic of group randomness. The earliest inspiration for this algorithm comes from the movement of point charges and Coulomb force in physics, as shown in. The genetic EM algorithm formulates criteria such as total force calculation between strokes, stroke power

calculation, stroke motion, local search, etc., so that each stroke in the population moves in the direction of the optimal stroke, which can be obtained after repeated several times. The optimal solution of the optimization problem can be used as an independent algorithm to solve the global optimization problem. It can also be integrated with other algorithms to form a new global optimization algorithm (global optimization algorithm, also known as modern heuristic algorithm, is an algorithm with global optimization performance, strong versatility and suitable for parallel processing) [8].

(2) Genetic EM algorithm classification. Because the genetic EM algorithm is a relatively new algorithm, domestic scholars have done very little research on the genetic EM algorithm, and the application of this method to the field of logistics distribution route optimization is still a blank. Fortunately, many researchers have done a lot of research on other optimization algorithms, including more advanced algorithms such as ant colony algorithm (it is a probabilistic algorithm used to find an optimal path, which comes from the behavior of ants finding paths in the process of looking for food) and genetic algorithm. Some scholars also apply these algorithms to the optimization problem selection of vehicle distribution routes, and have achieved certain results [9, 10].

(3) The principle of genetic EM algorithm. Genetic EM algorithm is an iterative optimization strategy. The genetic EM algorithm is affected by the idea of data loss. It was originally to solve the problem of parameter estimation in the case of missing data. The basis and convergence of the algorithm are effective. The basic idea is as follows: first, estimate the value of the model parameters based on the given observation data. Then estimate the value of the missing data based on the parameter value estimated in the previous step, and then add the missing data previously estimated based on the estimated missing data. Re-estimate the parameter values from the observed data, and iterate repeatedly until the final convergence, and the iteration ends. Taking the question in this article as an example, assuming that the logistics has to pass through k nodes, then there are $(k-1)!$ ways to move, so how to make the logistics journey the shortest?

First, Figure 1 randomly selects a point from the known feasible region as the initial stroke, calculate the objective function $G(x)$ of each stroke in a initial stroke, and calculate the optimal stroke $S^?$ of the objective function as the current optimal journey. Then set the local search parameters. Local search refers to an improved solution for searching the current population within a single travel range. According to the set local search parameters, calculate the various dimensions of the stroke, obtain the movement direction of the stroke and the maximum feasible step length, and perform a local search along this direction until a better
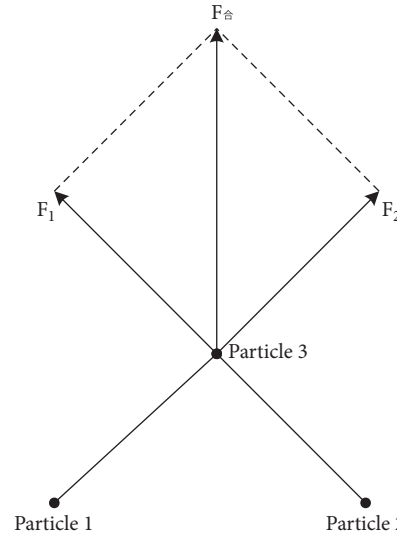
Figure 1: Schematic diagram of EM algorithm.

solution is found. After the local search is completed, the population is updated once, the objective function value of each trip after the population update is recalculated, and the optimal trip of the objective function is updated. Then calculate the charge of each stroke according to the value of the objective function of each stroke, and then calculate the resultant force of each stroke through the formula of simulating electromagnetic field force and the principle of superposition. In the process of calculating the resultant force, the value of the objective function of each stroke is greater than the value of the objective function of the other two strokes. In this case, the (good) stroke with a small objective function value absorbs the stroke with a large objective function value, and the stroke with a large objective function value (bad) rejects the stroke with a small objective function value. The final stroke moves in the direction of the resultant force, and the feasible step length of each step of the stroke can be calculated by the following formula. At this time, the position of each stroke is updated, and the iteration of similar electromagnetic algorithms is completed [11, 12].

2.2. Bayesian Network

(1) *Concept.* Bayesian network (BN), also known as causal probability network and belief network, is a network model for reasoning based on probabilistic uncertainty, as shown in. The basis of BN is Bayes' theorem (A theorem about the conditional probability of random events A and B, where $P(A|B)$ is the probability that A will occur if B occurs). BN is a network composed of nodes representing variables (including child nodes and parent nodes) and directed edges connecting these nodes. Random variables are represented by nodes, and the directed edges between each node represent nodes. The strength of the relationship is determined by the conditional probability. If there is no parent node, the information is represented by a priori

probability. BN is a directed graphic description based on the network structure. This is a graphic model used to express the connection probability of a set of variables. It is suitable for expressing and analyzing uncertain and probabilistic things. It cannot be incomplete or uncertain, and must be based on knowledge and information. BN is a directed acyclic graph (A directed acyclic graph refers to a directed graph without a loop. If there is a non-directed acyclic graph, and starting from point A to B via C, it can return to A, forming a ring. Changing the edge direction from C to A from A to C becomes a directed acyclic graph. The number of spanning trees of a directed acyclic graph is equal to the in-degree product of nodes with non-zero in-degree.), and the acyclic graph (DAG) is composed of variable nodes and directed edges connecting these nodes [13, 14].

BN is a mathematical model based on random reasoning, as shown in Figure 2. The emergence of BN is to solve the problem of imperfection and uncertainty. This method helps to solve the uncertainty and correlation of complex equipment, and has great benefits for reducing risks. Probability theory is the theoretical basis of BN and an accurate solution to problems in probability theory. The BN theoretical system often uses some basic probability formulas, such as conditional probability, probability law, total probability, Bayesian formula, etc. [15].

(2) *Features.* BN can qualitatively express complex causal or random relationships between events in a concise graphical way. With some prior information, these relationships can also be expressed quantitatively. The topology of the network is usually determined based on specific survey topics and questions. One of the current research hotspots of BN is how to automatically determine and optimize the network topology through learning [16].

(3) *Bayesian Formula.* In order to solve the joint probability problem, there are usually different calculation methods for different probability events. The following are several common probability calculation formulas, in which the probability value is represented by p.

(i) *Multiplication Formula.* If X and Y are two of the basic probability events of event Z, and the probability of any of these events is not 0, the calculation formula:

$$p(XY) = p(X)p(Y|X),\qquad(1)$$

$$p(XY) = p(Y)p(X|Y).\qquad(2)$$

(ii) *Conditional Probability.* If X and Y are two of the basic probability events of event Z, there is a conditional relationship between the occurrence of event X and event Y, and the probabilities affect each other. If event X occurs and the
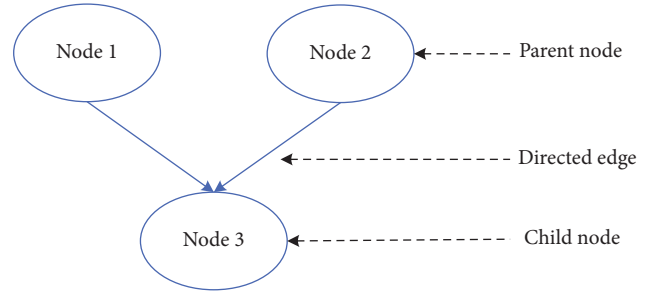
Figure 2: BN model.

probability is not 0, then the probability of event Y occurs:

$$p(Y|X) = \frac{p(XY)}{p(X)}.\qquad(3)$$

Similarly, if event Y occurs and the probability is not 0, then the probability of event X: Figure 2.

$$p(X|Y) = \frac{p(XY)}{p(Y)}.\qquad(4)$$

(iii) Using the multiplication formula and the conditional probability formula to generalize the total probability formula of the big event Z, assuming that $K_1, K_2, \ldots, K_n$ are the sub-events of event Z, and the probability of occurrence of the event is not 0, then the total probability formula of event Z can be expressed as:

$$p(Z) = \sum_{i=1}^{n} p(K_i)p(Z|K_i).\qquad(5)$$

From this, the conditional probability formula of the event K_i can be scored:

$$p(Z|K_i) = \frac{p(Z)}{\sum_{i=1}^{n} p(K_i)}.\qquad(6)$$

(iv) *Bayesian Formula.* Assuming that $K_1, K_2, \ldots, K_n$ are the sub-event groups of event Z, the probability is not 0, and the event X is any event among them, and the probability is not 0, then the above generalization can be obtained:

$$p(K_i|X) = \frac{p(K_i)p(X|K_i)}{\sum_{j=1}^{n} p(K_j)p(X|K_j)}.\qquad(7)$$

(4) *BN Mathematical Model.* BN modeling is a relatively complicated process, which requires patience and care. When modeling, the first thing to do is to determine the nodes, use the nodes to build the topology, then calculate the conditional probability distribution of each node, and finally test to determine the model. shows the basic model building process [17, 18].

(i) Selection of nodes. Bayesian network node variables are various elements of the emergency

system. It is very important to determine the node variables as comprehensively as possible for the accuracy of situation assessment. There are several main types of nodes in Bayesian networks. The first is the target node. This node represents the final problem that the model needs to solve. The second is the evidence node (also called the parent node), which is the most basic parent node in the network and has a value similar to the following: the prerequisites for Bayesian network inference are provided directly from the database or provided by expert knowledge. The third is to realize a component of the integrity of the Bayesian network. After determining the connected role nodes, the value of each node needs to be further determined. The value of each node is independent of each other and does not contain each other [19].

Assuming that there are m nodes in BN, and each node is represented by N, then: Figure 3.

$$P(N) = P(N_1, ..., N_m). \tag{8}$$

Taking m equal to 4 as an example, suppose there are four nodes H, I, J, and K, as shown in Figure 4:

According to the previous probability formula, we can get:

$$P(H, I, J, K) = P(K|J)|P(I|H, J)|P(J|H)|P(H). \tag{9}$$

(ii) The establishment of BN topology. The purpose of BN structure learning is to extract specific data and build a structure with a higher degree of fit. According to the previous knowledge, when the number of nodes is m, the number of structures can be expressed as a function:

$$G(m) = \sum_{i=1}^{m} (-1)^{i+1} \left[\frac{m!}{i!(m-1)!} \right] 2^{i(m-1)} g(m-i). \tag{10}$$

Using the scoring search method to perform statistics, R represents random variables, and S represents data collection, then:

$$R = \arg\max P(R|S), \tag{11}$$

$$P(R|S) = \frac{P(S|R)|P(R)}{P(S)}. \tag{12}$$

According to formula (12), the logarithm of both sides can be obtained

$$\text{Log}P(R|S)| = \text{Log}P(S|R)| + \text{Log}P(R). \tag{13}$$

Formula (13) is the score of R. In fact, $P(R)$ is usually evenly distributed, for which one can obtain

$$P(S|R) = \int P(S|R, \delta)|P(\delta|R)|d\delta. \tag{14}$$

In the formula, δ represents the size of the node corresponding to R.

When the data set S is independent and identically distributed, and R is the corresponding BN structure, the evaluation function for BD is

$$G_{BD}(R, S) = \text{Log}P(R) + \sum_{i=1}^{m} \sum_{j=1}^{h_i} \left[\text{Log} \frac{\Gamma(x_{i,j})}{\Gamma(x_{ij*} + y_{ij*})} + \sum_{c=1}^{d_i} \frac{\Gamma(x_{ijc} + y_{ijc})}{\Gamma(x_{ijc})} \right]. \tag{15}$$

When the data set S is independent and identically distributed, and R is a BN structure of y variables, the evaluation function for K2D is

$$G_{K2}(R, S) = \text{Log}P(R) + \sum_{i=1}^{m} \sum_{j=1}^{h_i} \left[\text{Log} \frac{(d_i - 1)!}{(y_{ij} + d_i - 1)!} + \sum_{c=1}^{d_i} \text{Log}(y_{ijc}) \right]. \tag{16}$$

If the BN structure information is uniformly distributed, the evaluation function for BDe is

$$G_{B\,De}(R, S) = \text{Log}P(R) + \sum_{i=1}^{m} \sum_{j=1}^{h_i} \left[Log \frac{\Gamma(x/h_i)}{(\Gamma(x/h_i) + y_{ij*})} + \sum_{c=1}^{d_i} \frac{\Gamma((x/d_i h_i) + y_{ijc})}{\Gamma(x/d_i h_i)} \right]. \tag{17}$$
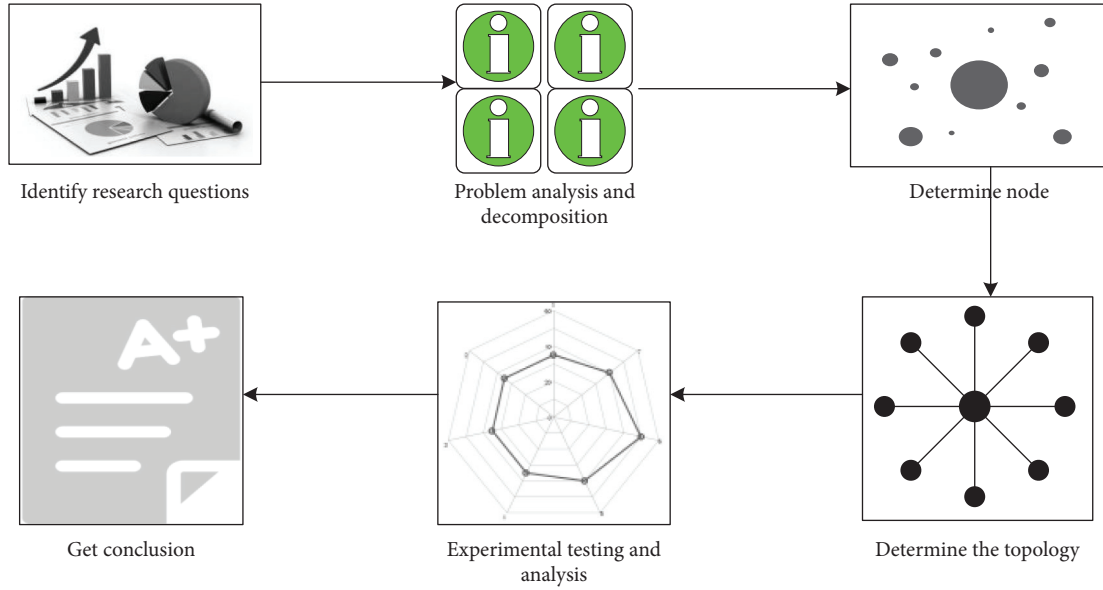
FIGURE 3: BN model establishment process.

H	δ
1	0.5
0	0.5

H	I	δ
1	1	0.3
1	0	0.7
0	1	0.6
0	0	0.4

J	K	δ
1	1	0.8
1	0	0.2
0	1	0.1
0	0	0.9

I	J	δ
1	1	0.75
1	0	0.25
0	1	0.2
0	0	0.8
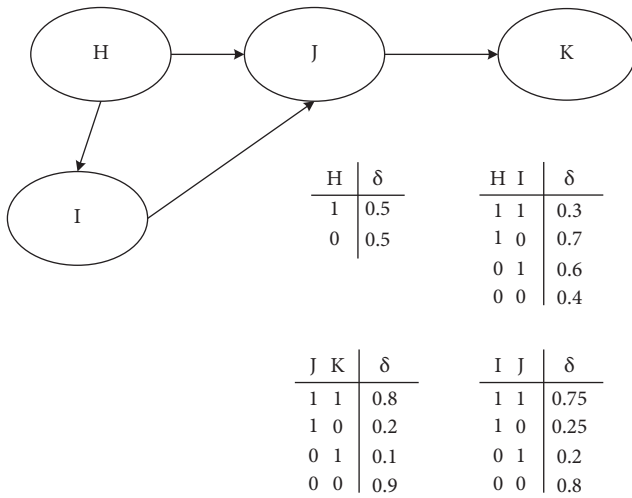
FIGURE 4: BN structure model of four nodes.

Or an evaluation function based on information theory, $LL(R|S)$ represents the logarithm of the likelihood of BN, then

$$G(R|S) = g(\delta)|R| - LL(R|S). \tag{18}$$

When $g(\delta) = 1$, there are

$$G_A(R|S) = |R| - LL(R|S). \tag{19}$$

When $g(\delta) = (1/2)\mathrm{Log}m$ represents the number of bytes

$$G_B(R|S) = \frac{1}{2}\mathrm{Log}|R| - LL(R|S). \qquad t \tag{20}$$

The selection of the evaluation function should be carried out according to the actual situation of the BN network, and the minimum evaluation function obtained by calculation can obtain the best parameters of BN [20, 21].

2.3. Plan to Control Logistics Delays

2.3.1. The Meaning of Logistics Delay.
The rapid development of e-commerce has brought more and more convenient online shopping services, and at the same time, it has also driven the rapid development of the logistics industry. shows the growth of online shopping users and the amount of online shopping transactions in the past five years.

It can be seen from Figure 5 that with the continuous development of e-commerce, the number of online shopping users has increased year by year, and the total amount of e-commerce transactions has also continued to increase. Due to the impact of the epidemic at the end of 2019, its online shopping transaction volume has declined, but online shopping users are still increasing, which also shows that during the epidemic, more and more people began to prefer online shopping. Figure 5.

However, there are also a lot of problems in the development process of the logistics industry. As a service-oriented industry, logistics is the most important thing to ensure the timely delivery of goods, so the problem of logistics delays is the most worthy of attention. Figure 6 shows the growth of the express delivery industry in recent years and the growth of the number of complaints about logistics delays by express users to the State Post Bureau.

It can be seen from Figure 6 that the number of complaints about logistics delays by express users to the State Post Bureau has been increasing year by year, and the number of express delivery delays accounted for more than 25% of the total number of complaints, such a serious delay has brought a lot of discomfort to e-commerce merchants, operators and users. For e-commerce merchants, delays in e-commerce logistics have brought them credibility and economic losses. Major e-commerce platforms regard logistics as an indicator for customers to evaluate their online stores. The poor experience of logistics delays will affect the reputation of the customer's store, damage the reputation of

(a)

(b)
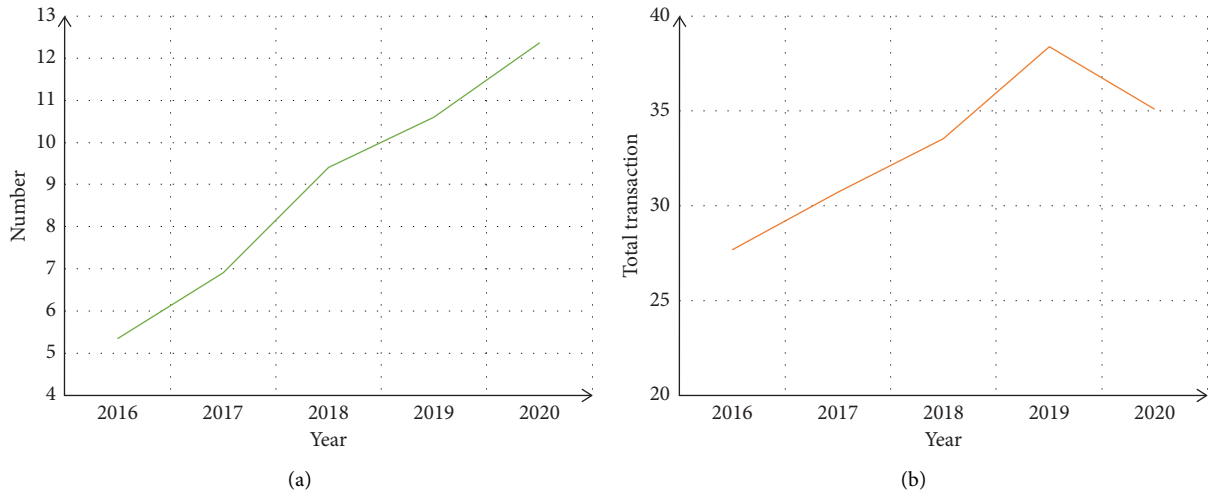
FIGURE 5: The development of online shopping and express delivery in China in recent years. (a) Growth of online shopping users. (b) Growth of online shopping transactions.

(a)

(b)
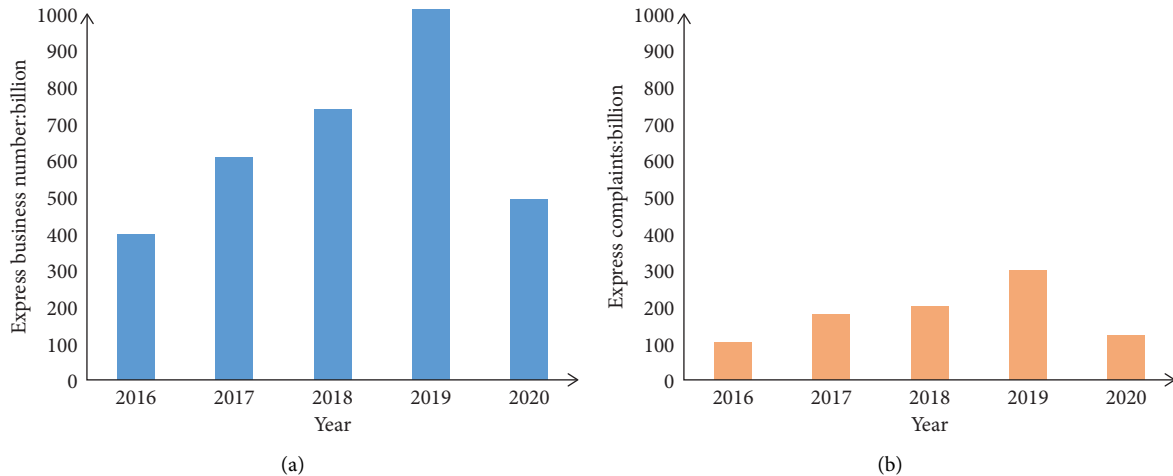
FIGURE 6: The number of express complaints about logistics delays in recent years. (a) Growth in express delivery business. (b) Delay in express delivery complaints.

the store, and affect the decision of other potential customers whether to purchase the store's goods [22, 23].

2.3.2. Reasons for Logistics Delays. According to the direct personnel contacting the goods, it can be divided into corporate factors (including logistics company factors, provider factors, courier factors), customer factors and other factors. Logistics company factors include arranging logistics drivers for loading, transportation schedules and other factors; providers include delivery speed, packing speed, etc.; the courier includes the speed of handover, telephone notification to customers, etc.; customer factors include pickup speed, fuzzy receiving address, incorrect phone number, etc.; other factors include road traffic conditions, weather factors, emergencies, and so on. Table 1 shows the statistics of the proportions of various factors, among which the corporate factor is the most important factor.

2.3.3. Features. When e-commerce logistics is delayed at the starting node, the delay will spread downstream through online shopping products. Affected by this, e-commerce logistics continues to lag downstream neighboring nodes, and there is a continuous delay between this point and downstream neighboring nodes. Affected by interference factors, there is a complicated correlation between the initial delay in e-commerce logistics at the starting node and the adjacent delay in downstream adjacent nodes. Initial delays in commercial logistics can increase, pass on, or weaken the impact of adjacent delays. Delivery companies have different factors in e-commerce logistics delay management methods, leading to the introduction of corporate and customer factors. The degree of correlation between successive delays in e-commerce logistics caused by various factors varies. Figure 7 is a schematic diagram of the spread of logistics delays [24].

TABLE 1: Proportion of reasons for logistics delays.

Factor	Include	Proportion (%)
Company	Logistics company	56.2
	Enterprise store	27.9
Customer	Pickup speed	11.6
Other	Natural factors	4.3
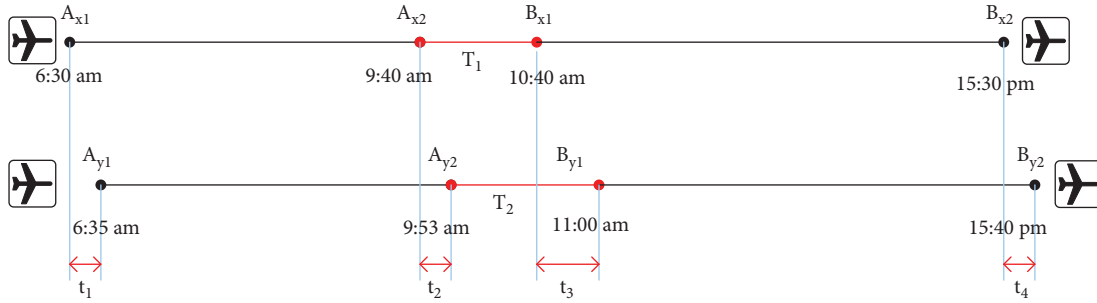
FIGURE 7: Schematic diagram of the spread of logistics delays.

2.3.4. Control Measures for Logistics Delays. Based on the analysis of the causes of logistics delays, it can be carried out from the aspects of enterprises, customers, etc., and other factors can be avoided if they can be avoided. On the corporate side, improvements or optimizations can be made in terms of employees, equipment, management methods, etc., while on the customer side, communication with them can be strengthened, and logistics awareness can be promoted. In other aspects, such as sudden traffic jams, if the transportation cost is similar, the detour is faster if it can take a detour. In severe weather, it can choose a better logistics method. ① In terms of corporate employees, improving the working level of the staff not only guarantees the work quality of the courier, but also ensures that the behavior of the customer during the delivery process meets the requirements of the courier, and avoids the delay caused by the following reasons in e-commerce logistics: operation error. Human factors are now an important reason for e-commerce logistics delays. Improving the work level of personnel is the key to avoiding unreasonable delays and an important means to control the spread of delays. To improve staff credibility, it is necessary to strengthen the training and management of express company employees. The staff of courier companies need to have professional knowledge and skills to effectively avoid delays caused by seal errors, misdelivery, and failures in timely response. At the same time, having professional knowledge and a sense of responsibility is an important guarantee for the courier staff to correctly guide the delivery behavior of customers and avoid problems such as blurred addresses and telephone errors. ② Stability of working equipment. The failure of logistics operation equipment is an important factor causing delays in e-commerce logistics. By improving the reliability of logistics equipment, it can effectively reduce the possibility of logistics delays caused by corporate factors in e-commerce. For this reason, this section focuses on the maintenance and repair management of logistics equipment. Working and deploying the correct logistics equipment are two aspects of improving equipment reliability. One is the

regular maintenance and overhaul management of logistics equipment. The logistics company's regular maintenance and overhaul management of logistics equipment can effectively avoid delays caused by mechanical equipment failures, increase equipment reliability, reduce potential delays in the cradle, and provide guarantee for e-commerce logistics operation planning. The second is to introduce appropriate logistics equipment. The introduction of suitable logistics equipment by logistics companies can effectively improve the efficiency of logistics operations, reduce the actual operation time of e-commerce logistics, and increase the "slack time" in the logistics operation process. If there is a delay in e-commerce logistics, e-commerce logistics is one of the nodes. The longer the "relaxation time" between the two, the more helpful it is to absorb the delay time and reduce the spread of the delay. ③ Reasonably arrange the priority of logistics business. Due to corporate factors, enterprises have poor management of e-commerce logistics delays. The spillover effect of delays is mainly manifested in transmission, and the delay absorption rate is very low, which reduces customer satisfaction with carriers and shipments, and damages the company's reputation. If e-commerce logistics is delayed due to corporate factors, the company is fully responsible, and the company must have greater control over the increasingly serious logistics delay. This article recommends that companies adopt new business priority allocation to solve the problem of e-commerce logistics delays caused by corporate factors. That is, change the logistics route and choose the route with higher priority [25].

3. EY Model Logistics Delay Control Simulation Experiment

3.1. Design of EY Model

3.1.1. Design Ideas. According to the number m of nodes passed by the logistics, the delay of each node logistics will affect the logistics delay of the two goods before and after, as

shown in Figure 8, the entire logistics business process of online shopping from receipt to delivery is connected. The goods pass through x_1, x_2, ..., x_i, ..., x_m logistics nodes in sequence according to the logistics plan. i is a positive integer. The delay of online shopping at specific logistics nodes will inevitably affect the timeliness of subsequent logistics nodes. In the process of logistics business, there is "logistics slack" in online shopping products. For example, an online shopping product will stay at a specific logistics node for an hour, but in reality, the staff of the logistics node will complete half of the logistics transfer activities such as weighing, loading and unloading, and product handling. The remaining 30 minutes are "logistics slack," but "logistics slack" may not be enough to absorb the delay time of the product operation on the previous node, leading to the spread of delays. If the user suggests to the carrier that the delivery of online shopping goods from node x_i to node $x_i + 1$ is delayed, the goods will continue to pass through nodes x_i, $x_i + 1$...x_m to the preliminary planned route of subsequent logistics business activities. Assuming that in the process of sending online shopping products from node x_i to node $x_i + 1$, the delay cannot be fully absorbed due to the interference factor "logistics slack," then the arrival of online shopping products at node $x_i + 1$ will be delayed; if $x_i + 1$ is affected by the delay time of the node after receiving, "logistics slack" cannot fully absorb the remaining delay time in the processing of online shopping products from node $x_i + 1$ to node $x_i + 2$, online shopping arrival delay node $x_i + 2$ products will certainly continue to be delayed. This paper mainly studies the delay propagation of online shopping products from node $x_i + 1$ to delivery node x_n. Node $x_i + 1$ is defined as the starting node and node $x_i + 2$ is the adjacent node; the delay of online shopping products at the starting node is defined as the initial delay. The delay of online shopping products on adjacent nodes affected by the initial delay is defined as the adjacent delay. The delay of online shopping products affected by the initial delay is defined as the delay of the delivery node. The delay time of the online shopping product at the starting node is defined as the initial delay time, that is, the difference between the time when the online shopping product actually arrives at the starting node and the planned time. The node delay time when the online shopping product is adjacent is defined as the adjacent delay time, that is, the difference between the actual time and the planned time for the online shopping product to reach the adjacent node; the delay time of the online shopping product at the delivery node is defined as the total delay time. This is expressed as the difference between the actual time and the planned time for the online shopping product to arrive at the delivery node. Online shopping products are affected by initial delays in the distribution process, and online shopping products continue to reach downstream nodes. Delay nodes (except the start node) are defined as downstream delay nodes [26, 27].

Figures 7 and 8 original delay model is improved and simplified, and the EY model is obtained. The model comparison is shown in Figure 9. The numbers in the figure represent various causes of logistics delays. Figure (a) is the model before improvement, and figure (b) is the model after improvement.

3.1.2. The Characteristics of the Improved Model.

The streamlined model is smaller in scale, faster in calculation, and suitable for processing large data sets. The improved model combines related factors, its structure is clearer, the calculation path is clearer, there is no redundant calculation, the simplified model is smaller in scale, faster in calculation, and suitable for processing large data sets. In the logistics delay forecast, the effect is better and the accuracy rate is high.

3.2. Simulation Experiment of Logistics Delay Control Based on EY Model.

This article simulates the simulation experiment of logistics transportation route optimization, and the selected EY model is feasible and efficient. Collecting data before the experiment. Before the experiment, this article investigated the complaints, data management, delayed mediation, problem handling, etc. from many domestic express company users. The data includes the number of courier services and delay factors for online shopping products. The selected courier tracking number is one of the courier's working group, indicating that the delayed item was purchased online, or the logistics information of the delayed item indicates that the seller shipped it. This article collected 11,000 e-commerce logistics delay data, of which 8,500 valid data. The number of e-commerce logistics delays caused by corporate factors was 6,213; the number of delays caused by customer factors was 2,101; and the number of e-commerce logistics delays caused by other factors was 186. As shown in Table 2.

The experiment selected 18 logistics nodes where a logistics company is about to transport goods. The purpose is to calculate a more reasonable logistics distribution route to make the transportation mileage the shortest. The logistics node is represented by two-dimensional coordinates (i, j), as shown in Table 3.

According to the permutation and combination algorithm, there are a total of 18 paths, including more than 6,400 trillion ways. It is basically difficult to calculate. Using the model optimization algorithm designed in this paper, first select a node as the center node, the coordinate center point is (100, 100), because considering that the node closest to the coordinate center point is the node number 15, the center node coordinate is selected as the number 15 (126, 85). According to the previous formula, take m as 18. After iterative calculation, the better mileage shown in Table 4 is obtained, and the corresponding comparison is shown in Figure 10.

The experimental results found that the calculation of the optimal route for the total mileage of the logistics transportation route calculated by the EY model has high accuracy, few iterations of the calculation, and high calculation efficiency. The optimal route obtained after 9 calculations is 1106 km. Among them, the number of occurrences of the optimal route is 4 times, ranking first. The longest calculation is 1225 km, the distance error is less than 1.1%, and the calculation accuracy is as high as 98%.
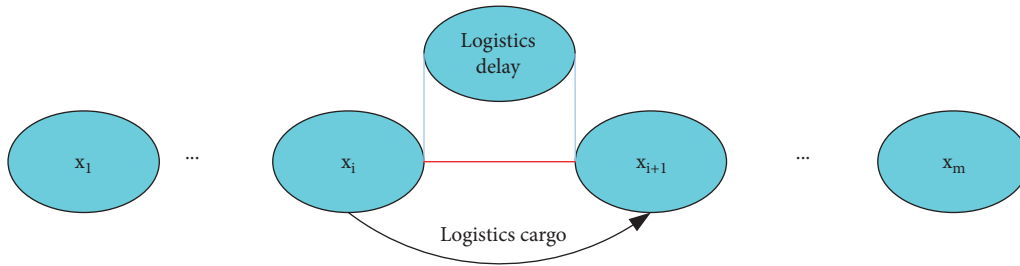
FIGURE 8: Schematic diagram of logistics delays.

(a)   (b)
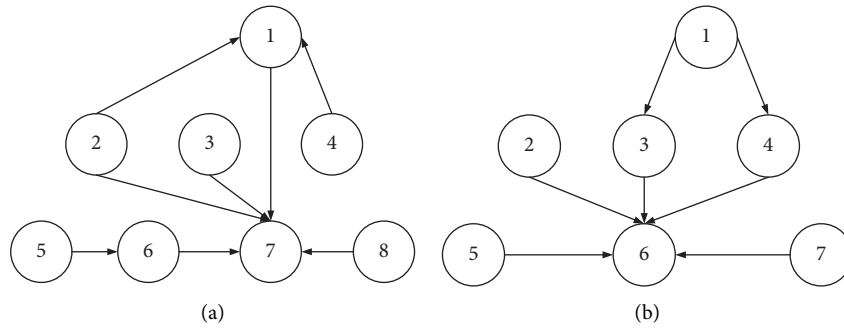
FIGURE 9: Logistics delay model.

TABLE 2: E-commerce logistics delay data.

Genre	Number	Proportion (%)
Valid data	8500	77.27
Business delays	6213	73.09
Customer delays	2101	24.72
Other delays	186	2.19
Logistics delay total	11000	

TABLE 3: Distribution of logistics nodes.

Node code	i	j
1	116	9
2	104	19
3	70	132
4	179	139
5	159	62
6	5	94
7	176	85
8	145	24
9	173	175
10	127	165
11	72	140
12	83	31
13	35	117
14	21	142
15	126	85
16	17	19
17	23	196
18	67	123

TABLE 4: Calculation results of EY model.

Total distance (km)	Number of calculations	Route
1225	2	A
1106	4	C
1139	3	B
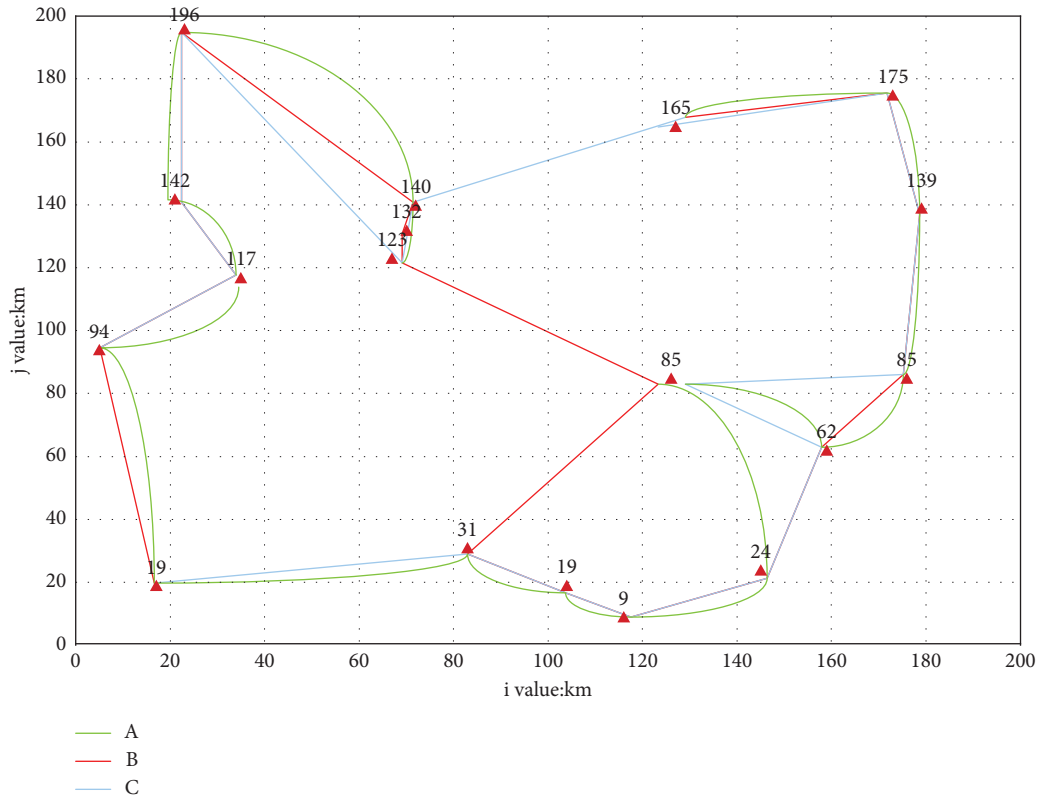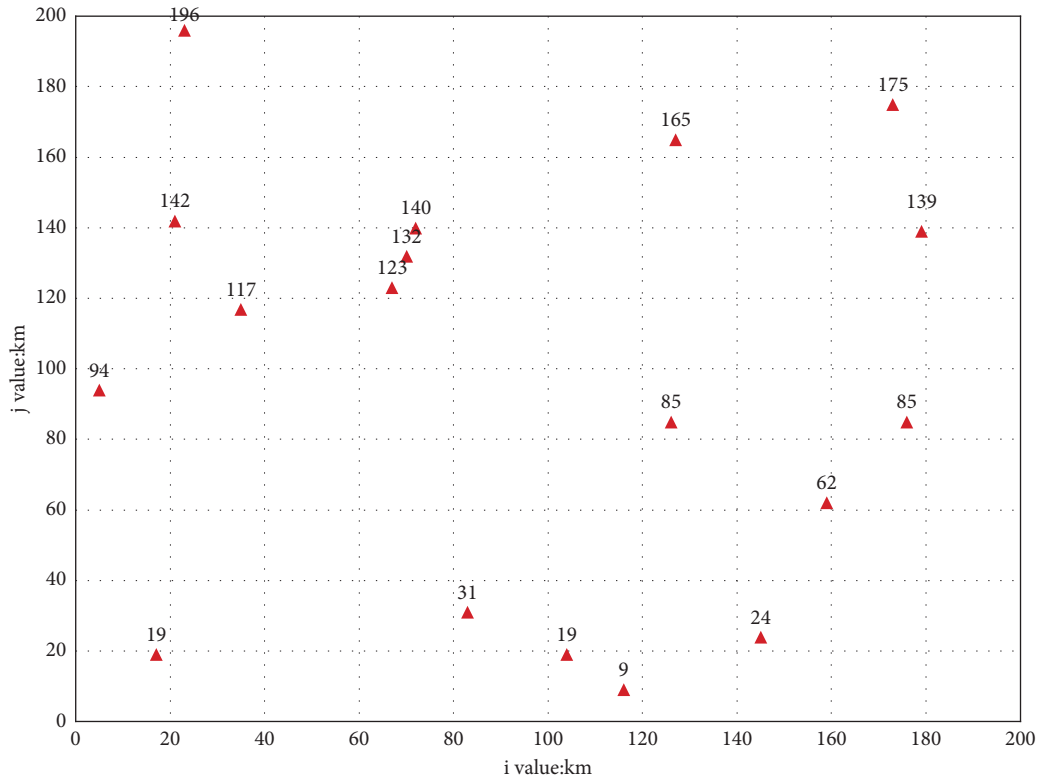
Figure 10: Results of route calculation by EY model. (a) Node distribution map and (b) Calculate a high-quality roadmap.

4. Discussion

The electromagnetic-like mechanism of logistics transportation path optimization algorithm simulates the mechanism of charged particles in physics, so that the particles move to the optimal solution. Particle turbulence allows the local search mechanism to search better both locally and globally. In order to obtain the best solution, the goal of preventing the best solution from slipping through the network has been achieved.

5. Conclusions

The EM algorithm proposed in this paper has made a certain breakthrough in solving the problem of logistics transportation route optimization. However, there is still room for further research on this issue. This article only conducted simulations with careful consideration of distance and calculation of increased vehicle restrictions and vehicle capacity restrictions. Road congestion is frequent in logistics and transportation, and short-distance routes will waste a lot of time due to congestion, resulting in reduced transportation efficiency. Therefore, in future research, it is necessary to increase congestion parameters and time parameters, comprehensively consider the length of the trip and the time of the trip, and obtain an algorithm that matches the actual logistics road conditions.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] S. Chaudhary, S. Indu, and S. Chaudhury, "Video-based road traffic monitoring and prediction using dynamic bayesian networks," *IET Intelligent Transport Systems*, vol. 12, no. 3, pp. 169–176, 2017.

[2] I. Chaturvedi, E. Ragusa, P. Gastaldo, R. Zunino, and E. Cambria, "Bayesian network based extreme learning machine for subjectivity detection," *Journal of the Franklin Institute*, vol. 355, no. 4, pp. 1780–1797, 2017.

[3] J. H. Bolt and L. C. van der Gaag, "Balanced sensitivity functions for tuning multi-dimensional Bayesian network classifiers," *International Journal of Approximate Reasoning*, vol. 80, pp. 361–376, 2017.

[4] X. F. Liang, H. D. Wang, H. Yi, and D. Li, "Warship reliability evaluation based on dynamic bayesian networks and numerical simulation," *Ocean Engineering*, vol. 136, pp. 129–140, 2017.

[5] Z. Zhang, F. Dong, and L. Xie, "Data-Driven Fault Prognosis Based on Incomplete Time Slice Dynamic Bayesian Network," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 239–244, 2018.

[6] R. Kawabe, H. Ito, H. Yamashita, and M. Goto, "Hierarchical Structure Learning in a Bayesian Network for the Analysis of Purchasing Behavior," *Total Quality Science*, vol. 4, no. 3, pp. 99–108, 2019.

[7] D. E. Jacob, N. E. Fenton, T. Noguchi, and D. Lagnado, "Resolving the so-called "probabilistic paradoxes in legal reasoning" with Bayesian networks," *Science & Justice: Journal of the Forensic Science Society*, vol. 59, no. 4, pp. 367–379, 2019.

[8] Y. Luo, D. Mcshan, D. Ray et al., "Development of a Fully Cross-Validated Bayesian Network Approach for Local Control Prediction in Lung Cancer," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 232–241, 2019.

[9] A. M. Mirhosseini, S. Adib Nazari, A. Maghsoud Pour, S. Etemadi Haghighi, and M. Zareh, "Probabilistic failure analysis of hot gas path in a heavy-duty gas turbine using Bayesian networks," *International Journal of System Assurance Engineering and Management*, vol. 10, no. 5, pp. 1173–1185, 2019.

[10] M. Yazdi, "A review paper to examine the validity of Bayesian network to build rational consensus in subjective probabilistic failure analysis," *International Journal of System Assurance Engineering and Management*, vol. 10, no. 1, pp. 1–18, 2019.

[11] H.-D. Wu and L. Han, "A novel reasoning model for credit investigation system based on Fuzzy Bayesian Network," *Procedia Computer Science*, vol. 183, no. 19, pp. 281–287, 2021.

[12] B. Dairy, J. Valverde-Rebaza, M. F. Moura, and A. D. A. Lopes, "A survey of the applications of Bayesian networks in agriculture," *Engineering Applications of Artificial Intelligence*, vol. 65, pp. 29–42, 2017.

[13] D. Ramos-López and A. R. Masegosa, "MAP inference in dynamic hybrid Bayesian networks," *Progress in Artificial Intelligence*, vol. 6, no. 2, pp. 133–144, 2017.

[14] S. Gheisari and M. R. Meybodi, "A new reasoning and learning model for Cognitive Wireless Sensor Networks based on Bayesian networks and learning automata cooperation," *Computer Networks*, vol. 124, pp. 11–26, 2017.

[15] S. S. Qian and C. A. Stow, "Comparative analysis of discretization methods in Bayesian networks," *Environmental Modelling & Software*, vol. 87, pp. 64–71, 2017.

[16] M. Javadi, G. Saeedi, and K. Shahriar, "Fuzzy Bayesian Network Model for Roof Fall Risk Analysis in Underground Coal Mines," *Journal of Applied Sciences*, vol. 17, no. 3, pp. 103–115, 2017.

[17] S. Kobayashi and S. Shirayama, "Selecting data adaptive learner from multiple deep learners using Bayesian networks," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4229–4241, 2021.

[18] B. Andrews, J. Ramsey, and G. F. Cooper, "Scoring Bayesian networks of mixed variables," *International Journal of Data Science & Analytics*, vol. 6, no. 1, pp. 1–16, 2018.

[19] K. Sachdeva, S. Aggarwal, and S. Aggarwal, "A Hybrid Approach for Neural Network in Pattern Storage," *Fusion: Practice and Applications*, vol. 6, no. No. 2, pp. 43–49, 2021.

[20] J. Lessan, L. Fu, and C. Wen, "A hybrid bayesian network model for predicting delays in train operations," *Computers & Industrial Engineering*, vol. 127, pp. 1214–1222, 2018.

[21] C. J. Butz, J. S. Oliveira, A. E. Dos Santos, and A. L. Madsen, "An empirical study of Bayesian network inference with

simple propagation," *International Journal of Approximate Reasoning*, vol. 92, pp. 198–211, 2018.

[22] A. Freire, M. Perkusich, R. Saraiva, H. Almeida, and A. Perkusich, "A Bayesian networks-based approach to assess and improve the teamwork quality of agile teams," *Information and Software Technology*, vol. 100, pp. 119–132, 2018.

[23] Z. Zsolt, W. Huang, S. Matthew, and L. Michael, "Multivariable and bayesian network analysis of outcome predictors in acute aneurysmal subarachnoid hemorrhage: review of a pure surgical series in the postinternational subarachnoid aneurysm trial era," *Operative Neurosurgery*, vol. 12, no. 6, pp. 603–610, 2018.

[24] F. Petitjean, W. Buntine, G. I. Webb, and N. Zaidi, "Accurate parameter estimation for Bayesian network classifiers using hierarchical Dirichlet processes," *Machine Learning*, vol. 107, no. 8–10, pp. 1303–1331, 2018.

[25] D.-M. K. Dennis, M. R. Williams, and M. E. Sigman, "Investigative probabilistic inferences of smokeless powder manufacturers utilizing a Bayesian network," *Forensic Chemistry*, vol. 3, pp. 41–51, 2017.

[26] Z. A. Abedallah and A. Ali, "Alwan, multi-objective chaotic butterfly optimization with deep neural network based sustainable healthcare management systems," *American Journal of Business and Operations Research*, vol. 4, pp. 39–48, 2021.

[27] A. Ramirez-Noriega, R. Juarez-Ramirez, and Y. Martinez-Ramirez, "Evaluation module based on bayesian networks to intelligent tutoring systems," *International Journal of Information Management*, vol. 37, no. 1, pp. 1488–1498, 2017.

Retraction

Retracted: Optimal Design of International Trade Logistics Based on Internet of Things Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Wang, "Optimal Design of International Trade Logistics Based on Internet of Things Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8781095, 11 pages, 2022.

Hindawi

Research Article

Optimal Design of International Trade Logistics Based on Internet of Things Technology

Yue Wang ⓘ

College of Economics and Management, Guizhou Institute of Technology, Guiyang 550003, Guizhou, China

Correspondence should be addressed to Yue Wang; kc3215@nyu.edu

With the global industrial integration, trade exchanges between countries are becoming more and more frequent and the competition for related international trade logistics is becoming more and more intense. This paper aims to study the optimization design of international trade logistics based on Internet of Things technology. For this reason, this paper proposes an RFID method, through which it is possible to master a journey in the actual international trade logistics process. And it optimizes and improves its related paths, shortens the time delayed on the road, and achieves the purpose of mentioning the carrying capacity. And for the actual use effect, the relevant experiments and analysis parts are designed to simulate it. The experimental results of this paper show that the improved and optimized international trade logistics shortens the carrying time by 28.6% and increases the carrying capacity by 34.2%, which greatly promotes the development of the relevant international trade logistics industry.

1. Introduction

The continuous development of the contemporary world economy toward integration has greatly promoted the development of international trade and international shipping, and logistics has played an increasingly powerful role in the development of the regional economy. After realizing the importance of logistics, all parts of the world are actively developing the logistics economy, and the competition among logistics is becoming more and more fierce, which makes the development of logistics more inclined to the specialization of functions, large-scale scale, modernization of technology, and scientific management. These changes make the competitive environment among logistics more cruel, but at the same time promote the rapid development of logistics in a better and more efficient direction.

Through the analysis of the construction of the existing international trade logistics platform, the development of international trade logistics has accelerated the operation of enterprise logistics, and it is also beneficial to the development of international trade. On the one hand, it can improve the loading and unloading processes, reduce the production cost, and improve the economic benefits of international trade,

which is beneficial to the long-term development of international trade; on the other hand, it enhances the expansion function of international trade, extends the basic function of international trade, and provides a new economic growth point for the development of international trade, thereby reducing the operational risk of international trade. From the perspective of international competition, the development of international trade logistics can help to establish an international trade logistics center and improve the comprehensive benefits of international trade logistics.

The innovation of this paper lies in the use of the key technology of the Internet of Things, that is, RFID radio frequency technology. And the technology is optimized so that the improved RFID technology can be better applied to the research topic of this paper. And the relevant national trade logistics is also analyzed and researched, and its memory is optimized and improved to a certain extent.

2. Related Work

With the development of economy and the expansion of globalization and internationalization, international trade logistics does not meet the basic functions of simple loading

and unloading, warehousing, and so on. How to adapt international trade logistics to the requirements of modern development, how to promote the sustainable development of economy and trade, and how to promote the regional development of surrounding areas have become one of the urgent problems to be solved. Lai M proposes a vehicle initial path optimization model considering uncertainty, considering vehicle capacity, customer time window, maximum travel distance, and road capacity [1]. The model proposed by Zhang et al. is formulated as a two-layer formula in which the upper layer determines the optimal choice of logistics infrastructure investment and green transportation mode subsidies to maximize the benefit-cost ratio of the entire logistics system [2]. Based on the relationship between suppliers, transit warehouses, and sellers in the collaborative logistics network, Xu et al. established a two-level programming model with random constraints considering the uncertainty of time. He proposed a genetic simulated annealing mixed intelligence algorithm to solve it. Numerical examples show that the method has strong robustness and convergence, and it can realize the rationalization and optimization of resource allocation in collaborative logistics network [3]. Aiming at the common distribution methods of urban logistics in the e-commerce environment, Yang considers the dynamic uncertainty of the urban road network based on the optimization of the distribution path [4]. Liu et al. proposed a stochastic model of postdisaster relief logistics to guide the tactical design of mobilizing relief material levels, planning initial helicopter deployments, and creating transport. He introduced a robust optimization method to deal with these uncertainties and derived a robust counterpart of the proposed stochastic model [5]. Rajasekar et al. reconsidered the problem of conflicting optimization goals, focusing on low inventory levels and short throughput times. It generates conversions of some performance metrics (such as delivery time and latency) to penalties. When manufacturing is the main economic activity of an industry, changes in production patterns pose challenges to the logistics system, forcing manufacturers to optimize the logistics system [6]. Dossou and Vermersch defined a study based on a cocreation approach that integrates all stakeholders (local authorities, companies, and citizens) to develop alternative solutions for road transport. It is used for the architecture and development of decision aids for simulating and optimizing alternative solutions to road traffic [7]. Malladi and Sowlatti proposed that biomass logistics was reviewed in a previous study aimed at classifying logistics operations, but the inherent issues and complexities and how they were incorporated into mathematical models were not discussed in detail. Their aim is to review the important features of biomass logistics operations, discuss how to incorporate them into mathematical optimization models, and explain new trends in biomass logistics optimization [8]. The above-mentioned literature is very good for the relevant logistics optimization ideas, and the relevant technical explanations are all in place. The proposed model is also studied in detail. However, the Internet of Things technology is not used in a certain combination, and the introduction of the relevant experimental process is not detailed enough.

3. International Trade Logistics Optimization Method

3.1. IoT Technology

3.1.1. Concept of Internet of Things. At present, there is still no authoritative definition of the Internet of Things. From the technical field, the Internet of Things refers to the integration of various technologies (sensors, wireless sensors, smart embedded, Internet, RFID, smart processing, nano, etc.), and through the Internet of Things, the interconnection of all things can be realized. It has the characteristics of precise control, reliable transmission, and comprehensive processing intelligence and perception. Figure 1 shows a typical IoT structure diagram [9].

3.1.2. IoT Architecture. The Internet of Things technology includes computer and Internet technology. It identifies and manages physical objects through modern information technology and uses wireless communication technology to achieve information exchange [10]. The Internet of Things has obvious advantages, the most notable feature of which is that it greatly reduces production costs while improving economic benefits. The Internet of Things is usually divided into five parts, namely, readers, electronic tags, Internet of Things name resolution servers, Internet of Things middleware, and Internet of Things information publishing services.

At present, the Internet of Things is at the forefront of technology, the technology is strict and complex, and the technology is not mature enough [11]. From the analysis of the current application requirements of the Internet of Things, the academic community generally divides the Internet of Things into three levels (1. perception layer; 2. network layer; and 3. application layer) as shown in Figure 2.

Perceptual recognition technology is the core technology of the Internet of Things [12]. The perception layer is mainly composed of various sensors and their gateways, which are mainly used to collect data. Its specific data indicators include various types of digital data, video data, audio, and identification [13]. The key technologies of IoT data collection include sensor technology, RFID technology, multimedia information collection technology, two-dimensional code technology, and real-time positioning technology. Taking people as an example, the nerve endings of human sensory eyes, ears, throat, nose, and skin are equivalent to the perception layer mentioned above. In the Internet of Things, identifying objects and collecting information sources is the core function of the perception layer [14].

The three compartments in the Internet of Things are independent of each other, but there is a close connection between the three compartments. In the introduction to the network layer, people are also taken as an example. The network layer can be compared to the human nerve center, which is responsible for transmitting and processing the digital information taken by the perception layer. The technical functions on each compartment complement each other and play different roles according to different
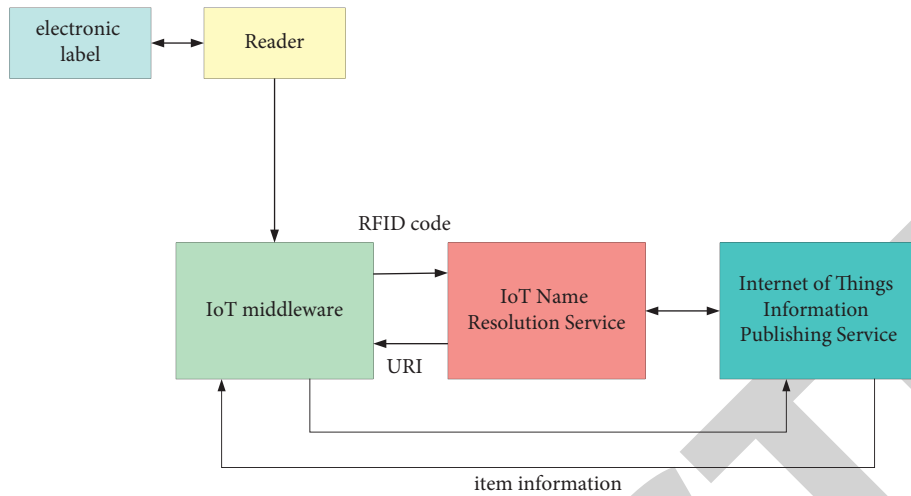
FIGURE 1: The structure of the Internet of Things.
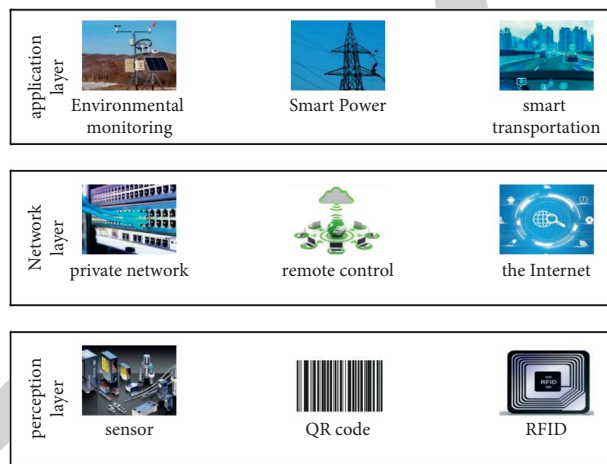
FIGURE 2: Internet of Things technology system framework.

environments to adapt to different application environments, thus forming a complete coping strategy for the compartment technology [15]. The security technology, identification and analysis, service quality management, and network management in the public technology are related to the above three layers of the IoT architecture to a certain extent, but they do not belong to a specific layer in the IoT technology.

3.1.3. Perception Layer Functions and Technologies. In view of the large amount of information obtained through perception functions, there is a large demand for geographic and spatial scope. The devices in the perception layer also need to form a self-organizing multinode network in a cooperative manner, and transmit data through the self-organizing network technology [16]. To realize the sensing function, the sensor network firstly touches and senses the sensed object in all directions through different sensors, and then comprehensively identifies and analyzes the acquired sensing data.

3.1.4. Key Technologies of the Perception Layer. RFID tags contain electronically stored information, so they can be identified within a few meters. Some tags do not need energy themselves and obtain energy from electromagnetic fields during identification; some tags themselves have power or energy, which can actively generate radio waves, the so-called electromagnetic fields tuned to radio frequencies. RFID technology has been used in most industries today [17]. Car manufacturers can know the progress of the car on the production line based on the label on the car, and the warehouse can also know the information of the goods in real time. In addition, radio frequency tags can also be applied to animals to facilitate the positive identification of animals. The radio frequency transponder installed on the car can be used to collect parking fees and also can be used to collect road tolls. Of course, this technology has the possibility of obtaining personal information without the consent of the individual, and this technology will infringe on the privacy of others to a certain extent [18].

The RFID application system is mainly used to control, detect, and track objects. The principle of RFID is not

complicated, its application system principle is actually very simple, the composition is also simple, and it is composed of two basic devices. The system consists of an interrogator, the so-called reader, and many transponders [19]. Its basic composition is shown in Figure 3.

Signal receivers also exist in RFID systems, which are often referred to as readers. The sophistication of readers varies widely based on the types of tags supported and the capabilities they perform. Its main function is to provide a data transmission method.

Receiving the radio frequency signal sent by the reader, the tag relies on the energy obtained by the induced current to send out the product information stored in the chip, or the tag can actively send a signal of a specific frequency. After reading and decoding the information, the reader sends the data and information to the central information system for further processing [20].

The signal transmitter exists in the RFID system, and its existing forms are also different according to the requirements of the application, and the label is its typical form. The tags include antennas, coils, memory, and low-power integrated circuits for control systems. The role of the label is like a barcode, and its role is to store information [21]. However, compared with barcodes, it has the characteristics of actively transmitting information [22]. The salient feature of RFID tags different from barcodes is that they do not need to be placed in the line of sight of the identifier, and RFID tags can also be embedded in the tracked object.

3.2. Power Transmission of the RFID Real-Time Positioning System.
In order to analyze the link characteristics and system indicators of the system, it is necessary to quantitatively analyze the power transmission mechanism of the RFID real-time positioning system. This includes the analysis of the path loss characteristics of the channel and the reflection characteristics of the tags. It first analyzes the path loss characteristics of the system.

The RFID system is a short-range wireless communication system, and the path loss of the system can be analyzed using the free space propagation model [23]. Assuming that the transmitting power of the reader is, it is transmitted outward through the antenna with a gain of. It can be seen from the formula that the relationship between path loss and distance is as follows:

$$P_{\text{channel}}|_{dB} = 20 \, \log\left(\frac{\lambda}{4\pi d}\right). \tag{1}$$

It can be known from the above formula that the maximum linear dimension D of the cross section of the transmitting antenna is related to the wavelength. Fraunhofer distance is defined as

$$d_f = \frac{2D^2}{\lambda}. \tag{2}$$

Let h_1 be the height of the reader, h_2 the height of the tag, and d the horizontal distance between the reader and the tag.

E_{tag} is the electric field strength received by the tag, and its expression is as follows:

$$\left|E_{tag}\right| = 2 \cdot \frac{\sqrt{30 \cdot \text{EIRP}}}{d} \cdot \, \sin\left(\frac{2\pi h_1 h_2}{\lambda d}\right). \tag{3}$$

Assuming that the EIRP is 4 W, the height of the reader and the tag are both 1 m, and the signal power received by the tag can be expressed as

$$P_{tag} = \frac{\left|E_{tag}\right|^2 \lambda^2}{480\pi^2},$$

$$P_{tag} = \sin^2\left(\frac{2\pi}{\lambda d}\right) \cdot \left(\frac{\lambda}{d\pi}\right)^2. \tag{4}$$

When the effective transmit power of the reader is 4 W, the received signal power obtained according to the free space attenuation model is as follows:

$$P'_{\text{tag}} = 4 \cdot \left(\frac{\lambda}{4\pi d}\right)^2,$$

$$P'_{\text{tag}} = \frac{1}{4}\left(\frac{\lambda}{\pi d}\right)^2. \tag{5}$$

As shown in Figure 4, the relationship between the received signal power of the tag and the distance obtained under the two models is given.

According to the working principle of the UHF RFID system, the energy received by the tag is roughly divided into two parts. One part is received by the tag antenna, enters the tag chip, and is rectified for the tag to work, and the other part is backscattered by the tag antenna [24]. To quantify the energy distribution between received and reflected, the tag antenna and tag chip can be modeled as shown in Figure 5.

v_s and I_s in the figure represent the peak voltage and peak current that can be induced by the tag antenna when it is not loaded; z_A represents the impedance of the tag antenna, z_L represents the impedance of the tag chip, and there are as follows:

$$Z_A = R_A + jX_A,$$
$$Z_L = R_L + jX_L. \tag{6}$$

Then, the reflection coefficient of the label is defined as

$$\Gamma = \frac{Z_L - Z_A^*}{Z_L + Z_A}. \tag{7}$$

Then, the power that can be obtained by the tag is as follows:

$$R_{\text{RF,in}} = \frac{1}{2}\text{Re}\left(V_S^* \cdot I_S\right). \tag{8}$$

Available power is as follows:

$$R_{\text{RF,in}} = P_{\text{avail}}\left(1 - |\Gamma|^2\right), \tag{9}$$

where
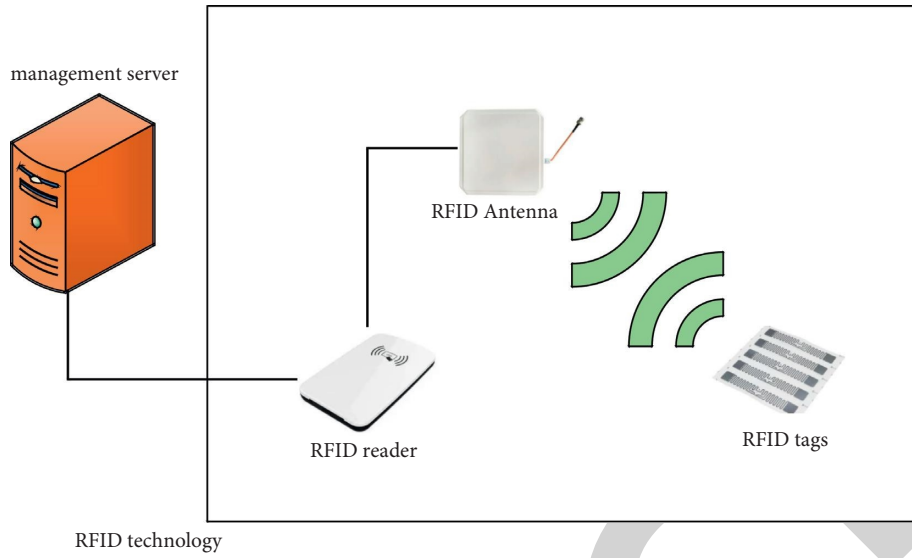
FIGURE 3: Basic composition of RFID.
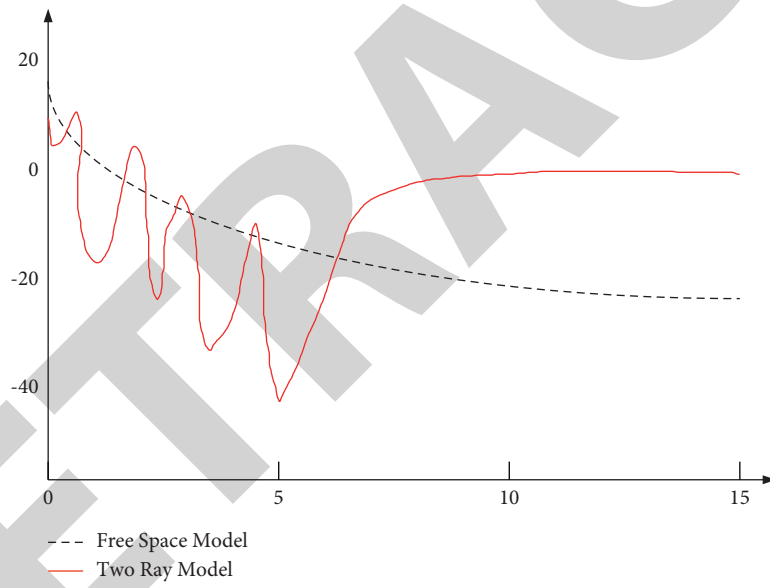
FIGURE 4: Relationship between received signal power and distance under the two models.
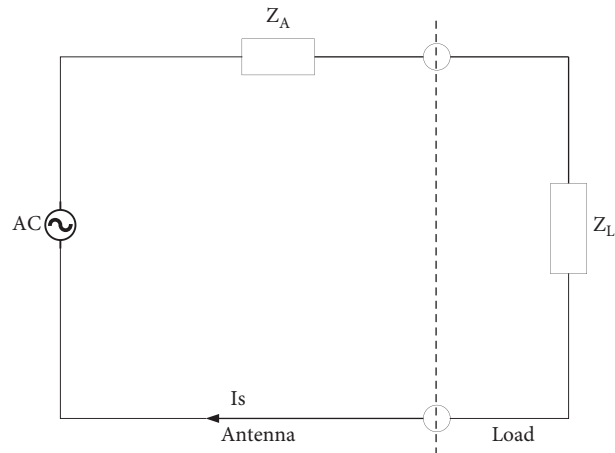
FIGURE 5: Label equivalent circuit.

$$P_{\text{avail}} = \frac{V_S^2}{8R_A}. \tag{10}$$

The load impedance of the tag during normal operation switches between two impedances is as follows:

$$
\begin{aligned}
Z_1 &= R_1 + jX_1, \\
Z_2 &= R_2 + jX_2.
\end{aligned}
\tag{11}
$$

The reflection coefficient and the power received by the tag also vary between the two states with

$$R_{\text{RF,in},1,2} = P_{\text{avail}}\left(1 - |\Gamma_{1,2}|^2\right). \tag{12}$$

Assuming that the probabilities of the two states appearing are $P1$ and $P2$, respectively, the average available power received by the tag is as follows:

$$P_{\text{RF,in,ave}} = p_1 P_{\text{RF,in},1} + p_2 P_{\text{RF,in},2}. \tag{13}$$

In order to calculate the power reflected back by the tag antenna, the resistance of the antenna can be divided into two parts: radiation resistance R_{rad} and lossy resistance R_{loss}, namely,

$$
\begin{aligned}
Z_A &= R_A + jX_A, \\
Z_A &= R_{\text{loss}} + R_{\text{rad}} + jX_A.
\end{aligned}
\tag{14}
$$

The current flowing through the antenna as the tag changes between the two states is as follows:

$$i_{1,2} = \frac{V_S}{Z_A + Z_{1,2}}. \tag{15}$$

Suppose the two states are equally likely to occur, both are as follows:

$$p_1 = p_2 = 0.5. \tag{16}$$

Then, the signal power reflected by the tag is as follows:

$$P_{bs} = \frac{P_{\text{avail}}}{4L_A}|\Gamma_1 - \Gamma_2|^2. \tag{17}$$

3.3. International Trade and International Logistics.

With the development of international trade, international logistics has been continuously improved. As a bridge of economic exchanges between countries, international logistics and international trade together constitute two indispensable aspects of international economic development [25].

Generally speaking, international logistics mainly face three barriers (Market and Competition, Financial Barriers, and Distribution Channels), as shown in Figure 6. The development of international logistics must balance the relationship between the actual cost of overcoming these barriers and the potential benefits of international trade, in order to obtain actual benefits through successful international operations.

Market and competitive barriers: market and competition barriers in international logistics mainly refer to market entry restrictions, information symmetry, pricing, and competitive environment [26].

First, market entry restrictions often create market entry barriers for imported goods through legislation or judicial practice.

Second, information symmetry is another barrier to international logistics. In addition to information asymmetry about market size, population, and competition conditions, the information used to clarify import operations and related documents is often not coordinated [27].

4. RFID Experiment Based on IoT

4.1. Distribution and Transportation Experiment.

In the international trade logistics, the logistics distribution center plays a linking role. Its upstream is international processing enterprises, and its downstream is major users. The international processing enterprise sends the processed products to the warehouse of the distribution center, and the distribution center completes the distribution of the processed products through transportation vehicles according to the requirements of customers. Because of the particularity of international trade logistics, enterprises are required to maintain a reasonable placement of finished products in a series of processes (warehousing and warehousing, loading and unloading, and distribution and transportation). The data collected during the distribution and transportation process and the corresponding equipment are shown in Table 1:

4.2. Example Application Experiment.

The calculation of genetic algorithm can be performed by the computer to verify its effectiveness in the optimization of cold chain logistics distribution path. With the help of Matlab 2011a, by writing the corresponding computer program of the model, the genetic algorithm is used to calculate the model, and the optimal solution of the problem is obtained. Taking city A as an example, the location coordinates of the distribution center and 10 customers, and the product demand of each customer are shown in Table 2.

At the same time, obtain the relevant receipt time window, unloading time, and stopover time in the international trade logistics center for statistics, as shown in Table 3.

Assuming the parameters corresponding to the traffic conditions in city A, the transport vehicles equipped in the distribution center are changed to Dongfeng Duolika transport vehicles (with a load capacity of 7 tons). Statistics are made on the number of transport vehicles required by these 10 distribution points. And taking into account the location coordinates of 10 customers, the demand for chilled meat, and the 7t load of refrigerated transport vehicles, the 10 customers are divided into the following 3 routes, as shown in Table 4.

Using Matlab 2011a calculation, decode the calculated solution space to get the optimal distribution route scheme. The results are shown in Table 5.
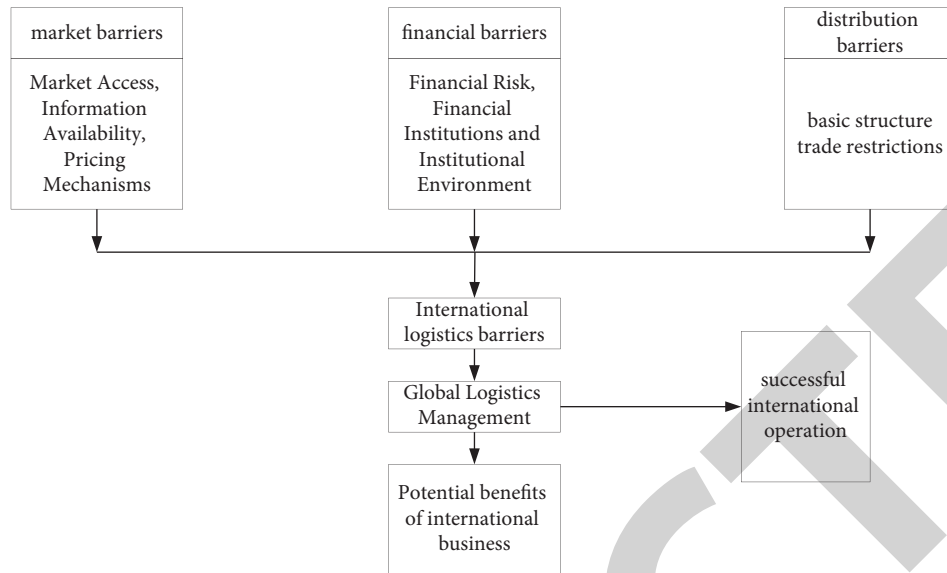
FIGURE 6: International logistics barriers.

TABLE 1: Collected data and equipment used.

Numbering	Link	Collected data	Equipment
1	Discharge	Distribution vehicle situation, cargo situation, distribution center information	RFID reader, measuring instrument information
2	Sales and processing	Workshop information, operating hours, operators, packaging information	RFID reader, video capture
3	Sales	Product information	RFID reader, barcode

TABLE 2: Statistical table of customer coordinates and product demand.

Serial number	X coordinate	Y coordinate	Demand
0	13426.14	2893.12	0
1	13462.11	2894.16	1
2	13428.16	2897.36	1.2
3	13422.62	2894.19	2
4	13429.18	2894.43	2
5	13481.64	2896.36	3.5
6	13492.11	2891.18	2
7	13466.23	2899.48	1
8	13428.36	2894.18	0.5
9	13492.16	2895.77	2.5
10	13411.12	2894.61	1.5

TABLE 3: Receipt time window, unloading time, and dwell time statistics.

Serial number	Earliest pickup time	Latest pickup time	Dwell time
0	6:00	17:30	0
1	7:15	14:30	0.5
2	6:00	15:30	0.75
3	7:00	16:00	0.25
4	7:00	10:30	0.25
5	6:30	8:00	0.75
6	6:45	13:00	0.5
7	6:45	10:30	0.5
8	6:30	17:15	0.33
9	6:00	12:00	0.5
10	6:00	15:00	0.25

TABLE 4: Route division table.

Vehicle serial number	Customer point	Number of customers	Vehicle capacity
1	1, 2, 4, 7, 8	5	6
2	5, 9	2	6
3	3, 6, 10	3	5.5

TABLE 5: Optimal route delivery table.

Vehicle serial number	Shipping order	Number of customers	Delivery mileage	Waiting time	Delay time
1	0-7-2-8-4-1-0	5	13.47	0	0
2	0-5-9-0	2	21.09	0	0
3	0-9-30-26-23-0	3	17.82	0	0

From the improved optimal route distribution table, we can see that after the improvement, the delay time in the logistics distribution process is basically 0, which can save time in the transportation process.

5. RFID Tags and Optimization Analysis

5.1. Modulation Mode Simulation. Different modulation methods can be achieved by changing the load impedance to change the current flowing through the antenna and the reflection coefficient of the antenna: ASK modulation can be achieved by changing the amplitude of the current, and PSK modulation can be achieved by changing the phase of the current. Under different modulation modes, the reflection coefficient of the tag is different, and the power entering the tag and the power reflected by the tag are different. Figure 7 shows the Smith chart of reflection coefficients under different modulation modes.

Comparing the above four modulation methods, we can see the following:

(1) Under ASK1 modulation, the available power entering the tag and the power reflected by the tag remain unchanged, which are 50% and 25%, respectively; under the other three modulation modes, the available power entering the tag decreases with the increase of the reflection coefficient.

(2) PSK and ASK2 have the same tag useable power (83%) and reflected power (17%), but under ASK2 modulation, part of the tag useable power is consumed by extra resistors and cannot be used effectively.

(3) ASK and PSK use nonlossy components to change the state of impedance, similar to PSK modulation. In the ideal case, the sum of the tag's available power and the tag's reflected power is a constant value. The tags of UHF RFID systems usually use load capacitance modulation to achieve ASK and PSK modulation from the tag to the reader. At this time, the power reflected by the tag is related to the modulation index and antenna loss, generally accounting for 10% to 20% of the total received power.

5.2. Tag Antenna Performance. By understanding the design of the tag antenna, we can see that the reduction or increase of the resonant frequency of the tag antenna can be achieved by adding capacitive or inductive loads. A capacitive load (low frequency) and an inductive load (high frequency) work together so that the tag antenna has two resonance modes, high and low. And the bandwidth of the S-parameter curve of the tag antenna can be changed by adjusting the length of the T-shaped microstrip structure and adjusting the depth of the U-shaped hollow structure. RFID tag antennas for metal backgrounds were designed and simulated using HFSS as shown in Figure 8.

It can be seen from the simulation results of the tag antenna that the two lowest points of the S-parameter curve of this tag antenna are 870 MHz and 915 MHz, respectively, and these two frequencies are also the resonant frequencies of the tag antenna.

The comparison between the measurement results of the tag antenna and the simulated junction impedance curve is shown in Figure 9. At 915 MHz, the antenna impedance obtained by simulation is (13.5~120j) ohm, and the antenna impedance is (15.3~123j) ohm obtained from the test, and the tag antenna and the tag chip are well matched.

5.3. Optimization Results. In order to have a general understanding of the optimization results of international trade logistics based on the Internet of Things technology proposed in this paper, this paper makes a rough simulation of the trade logistics of city A and analyzes it. It analyzes and compares the improvement effect of this paper with the effect before improvement and makes statistics on five different logistics distribution centers. It counts the amount of goods carried in the same time period to reflect the optimization effect of trade logistics. The specific situation is shown in Figure 10.

Through the above comparative analysis, we can see that after the improvement, the logistics has a relatively obvious increase in the carrying capacity. Before the improvement, the carrying capacity of site 1 was 15,000 tons, accounting for 12.93%; the carrying capacity of site 2 was 19,000 tons, accounting for 16.38%; the carrying capacity of site 3 was 26,000 tons, accounting for 22.41%;
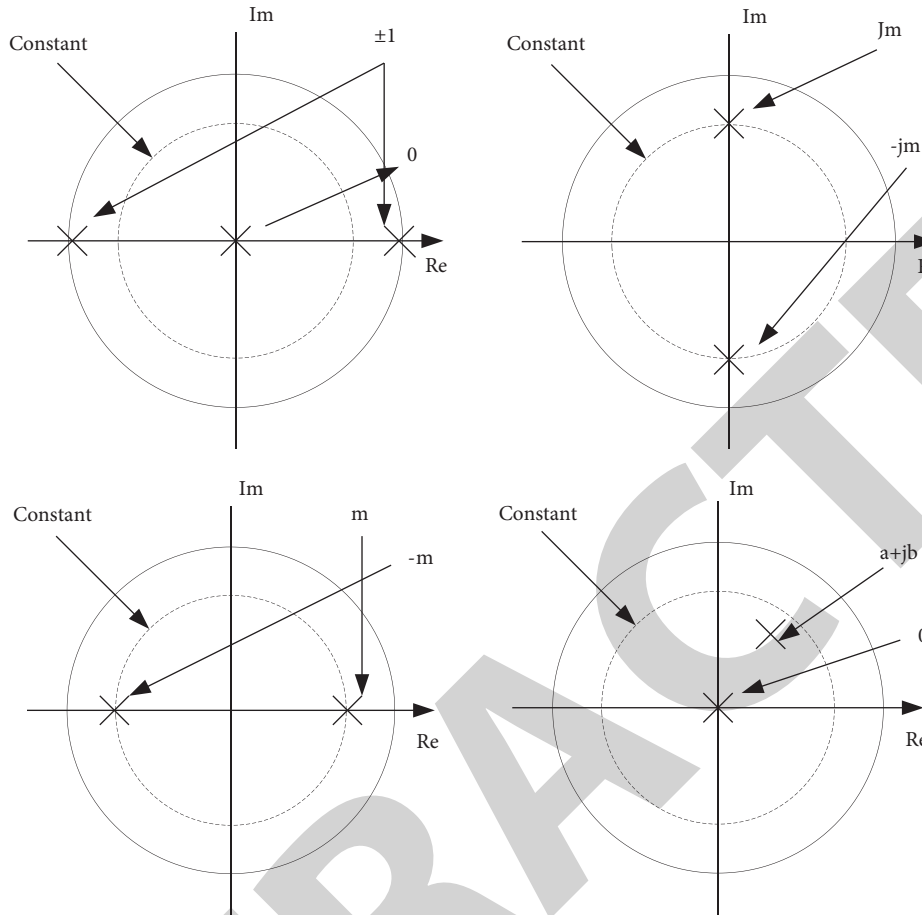
FIGURE 7: Reflection coefficients of labels under four different modulation modes.
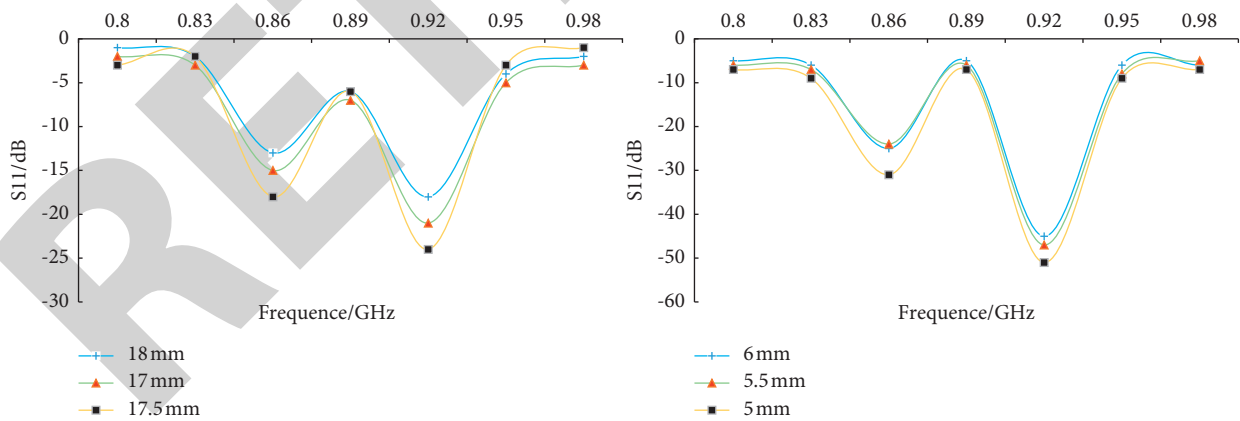
FIGURE 8: The frequency correspondence between the T-shape and U-shape of the S-parameter curve of the tag antenna.

the carrying capacity of site 4 is 31,000 tons, accounting for 26.72%; and the carrying capacity of site 5 is 25,000 tons, accounting for 21.55%. On the other hand, for the improved carrying capacity, the carrying capacity of site 1 is 25,000 tons, accounting for 15.82%; the carrying capacity of site 2 is 29,000 tons, accounting for 18.35%; the carrying capacity of site 3 is 34,000 tons, accounting for

21.52%; the carrying capacity of site 4 is 38,000 tons, accounting for 24.05%; and the carrying capacity of site 5 is 32,000 tons, accounting for 20.25%.

Based on the above analysis, we can conclude that the improved international trade logistics has increased the carrying capacity by 34.2% and shortened the carrying time by 28.6%, which has been significantly improved.
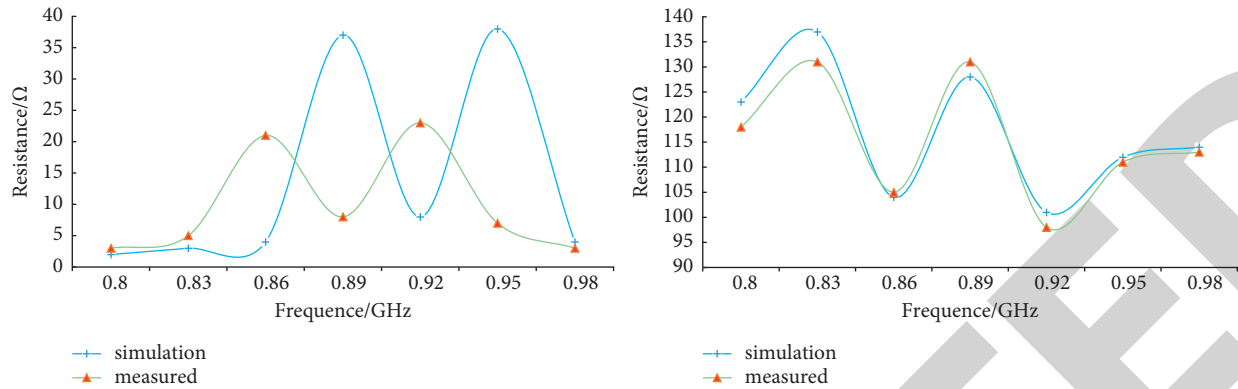
FIGURE 9: Comparison of measurement results of tag antenna resistance and reactance with simulation results.
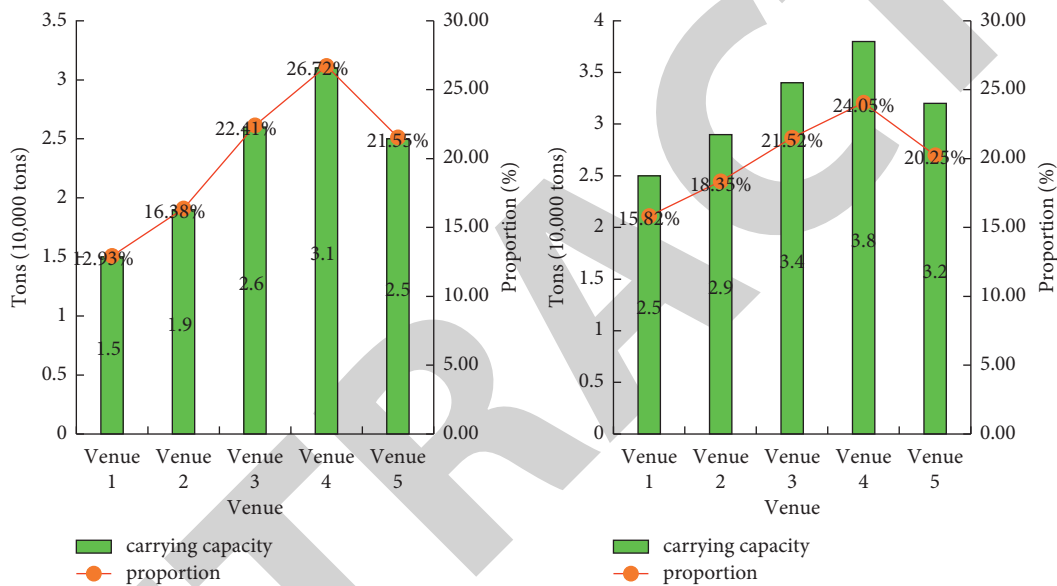
FIGURE 10: Comparison of logistics carrying capacity before and after optimization.

6. Conclusions

This paper mainly studies the optimization design of international trade logistics based on the Internet of Things technology, through the optimization and improvement of RFID technology, and the key technology of the Internet of Things. Through its tracking and identification of a path in the process of trade logistics transportation, and optimizing it, the optimal transportation path can be achieved in the actual transportation process. This also fully guarantees the international trade carrying capacity, carrying capacity, and carrying efficiency. And in the experimental part, a simulation experiment is carried out on the optimization of its logistics, the results also show that the scheme is feasible. And in the analysis part, the relative RFID technology is compared and analyzed to explore the actual effect of its optimization.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] M. Lai, H. Yang, S. Yang, and Z. Junhua, "Cyber-physical logistics system-based vehicle routing optimization," *Journal of Industrial and Management Optimization*, vol. 10, no. 3, pp. 701–715, 2017.

[2] D. Zhang, Q. Zhan, Y. Chen, and S. Li, "Joint optimization of logistics infrastructure investments and subsidies in a regional logistics network with CO2 emission reduction targets," *Transportation Research Part D: Transport and Environment*, vol. 60, no. 5, pp. 174–190, 2018.

[3] X. F. Xu, W. H. Chang, and J. Liu, "Resource allocation optimization model of collaborative logistics network based on bilevel programming," *Scientific Programming*, vol. 2017, no. 1, Article ID 4587098, 8 pages, 2017.

[4] S. Yang, "Optimization of urban logistics distribution path under dynamic traffic network," *International Core Journal of Engineering*, vol. 6, no. 1, pp. 243–248, 2020.

Research Article

A Hybrid Model for Leaf Diseases Classification Based on the Modified Deep Transfer Learning and Ensemble Approach for Agricultural AIoT-Based Monitoring

Maryam Saberi Anari [ID]

Department of Computer Engineering, Technical and Vocational University (TVU), Tehran, Iran

Correspondence should be addressed to Maryam Saberi Anari; saberi-m@tvu.ac.ir

As possible diseases develop on plant leaves, classification is constantly hampered by obstacles such as overfitting and low accuracy. To distinguish healthy products from defective ones, the agricultural industry requires precise and error-free analysis. Deep convolutional neural networks are an efficient model of autonomous feature extraction that has been shown to be fairly effective for detection and classification tasks. However, deep convolutional neural networks often require a large amount of training data, cannot be translated, and need a number of parameters to be specified and tweaked. This paper proposes a highly effective structure that can be applied to classifying multiple leaf diseases of plants and fruits during the feature extraction step. It uses a deep transfer learning model that has been modified to serve this purpose. In summary, we use model engineering (ME) to extract features. Multiple support vector machine (SVM) models are employed to enhance feature discrimination and processing speed. The kernel parameters of the radial basis function (RBF) are determined based on the selected model in the training step. PlantVillage and UCI datasets were used to analyze six leaf image sets containing healthy and diseased leaves of apple, corn, cotton, grape, pepper, and rice. The classification process resulted in approximately 90,000 images. During the experimental implementation phase, the results show the potential of a powerful model in classification operations, which will be beneficial for a variety of future leaf disease diagnostic applications for the agricultural industry.

1. Introduction

Before they reach maturity, the diseases in the leaves of fruits, citrus, wheat, and rice can have a significant impact on their yields [1]. This necessitates rapid and precise diagnosis of leaf diseases in fruits, citrus, wheat, and rice as well as early delivery of tailored cure [2]. Human investigation-based recognition of leaf diseases is severely hampered due to the huge area, and a model capable of tackling this issue is urgently required. The plant itself typically has its own recognition model when a leaf disease of a certain sort is recognized. A problem with storage capacity will occur if each disease of fruits, citrus, wheat, and rice has its own recognition model. Multitask learning allows features to be shared and aided by one another. This allows the leaf disease recognition models to meet current needs while reducing

storage space and increasing recognition accuracy. For this reason, a model that can detect and identify leaf diseases in fruits, citrus, wheat, and rice must be developed.

For image classification, the adoption of deep learning (DL) has been driven by recent advancements in computer vision and artificial intelligence [3–5]. Convolutional neural networks (CNNs) are an efficient model of DL structure. Self-learning, adaptability, and generalization are the hallmarks of CNN. The leaf diseases in fruits, citrus, wheat, and rice have been classified by machine vision approaches in numerous studies. Azadbakht et al. [6] achieved 82 percent correct classification of the normal and diseased rice leaves. Azadbakht et al. [6] employed machine learning (ML) approaches to analyze the wheat leaf hyperspectral data. Plant leaves are usually the first place where plant diseases are classified, and signs of most diseases may arise on the leaves

[7, 8]. CNN requires no manual feature extraction as in classic machine vision (MV) procedures, which rely on manual classification. Instead, CNN only requires visual input. In recent years, deep learning approaches, specifically CNNs, have become the preferred method to overcome learning challenges [9, 10].

The network's ability to self-learn allows it to complete the image classification process successfully [11]. It has become increasingly common in recent years to use DL image classification strategy in agriculture for detecting plant diseases [12–15].

Three typical rice leaf diseases were categorized using a DL network that was trained using DenseNet and Inception models with ImageNet pretrained. The accuracy percentage was found to be 94.07 percent [16].

Support vector machines (SVM) and DL were utilized by Jiang et al. [17] to identify four distinct rice leaf diseases. Using SVM and CNN, they were able to extract the features from various leaf images and recognize the extracted features, with an accuracy of 96%. Transfer learning-based powdery mildew disease detection was proposed by Shin et al. [18]. Using six different models, they compared their categorization findings. In this case, they found that ResNet-50 had satisfactory accuracy in identifying objects, while AlexNet had the quickest recognition time. In order to diagnose grapevine leafroll disease, Gao et al. [19] employed ML techniques to analyze hyperspectral images of grape leaves. The disease in grapevine leafroll can be diagnosed in its asymptomatic stage using this model, which achieved an accuracy of 89.93 percent [19]. DL networks were utilized by Long et al. [20] to recognize four different forms of camellia diseases based on image processing analysis. They employed ImageNet to train AlexNet and then built a novel fully linked layer with an accuracy of 91.25%. The pretrained model can greatly enhance accuracy when there are few samples available. Generally, we can utilize the same machine learning model to do many tasks using multitask learning, which is a common practice. For the strawberry Verticillium wilt classification, researchers applied the multitask learning network of the attention system. Additionally, 99.95% of the time, they properly diagnosed strawberry Verticillium wilt [21]. Plant phenotyping was studied by Dobrescu et al. [22] using multitask learning. Genotype classification, leaf counting, and PLA estimation were achieved using the ResNet-50 model trained on ImageNet. Each of the three locations has a significant advantage over the other options.

To simulate realistic situations of leaf disease, Arsenovic et al. [23] used images taken in a range of weather conditions, at various angles and at different times of day, all with varying backdrops. In their study, the authors propose a novel two-stage neural network (NN) approach for classifying plant diseases in the context of a real environment. Their model's accuracy was 93.67%.

Karthik et al. [24] presented a design based on residual learning to classify significant aspects for detecting infection in tomato leaves. Their proposed approach makes use of the PlantVillage dataset, which contains three diseases, and

CNN-learned features. They achieved 98 percent accuracy on the validation sets after fivefold cross-validation.

Sharma et al. [25] developed an approach for CNN model training. When tested on unlabeled data, the S-CNN model trained employing segmented images outperformed the F-CNN model by more than doubling its performance to 98.6% accuracy.

Sambasivam and Opiyo [26] suggested an approach for detecting cassava leaf disease in their study. They achieved an accuracy score of over 93% by using deep CNN built from scratch, class weight, and the Synthetic Minority Oversampling Technique.

Ma et al. [27] claimed that they used VGG to identify and classify four diseases in cucumber leaf leaves which resulted in crop loss. The classification was done on a vast volume of plant leaf images with the goal of classifying the presence or causes of disease in the leaves.

Too et al. [28] used four transfer learning (TL) models to classify the absence and presence of disease in plant leafs. These components included the VGG16, early V4, ResNet, and DenseNet structures. DenseNet, when compared to the other forms, generates more relevant responses with less processing time.

He also evaluated 12 different species of plants in another study [29], which were used to classify the presence or absence of disease in plants.

Kaya et al. [30] investigated datasets such as Flavia, Swedish, and UCI Leaf using TL for deep NNs in plant classification procedures. Similarly, Singh et al. [31] investigated the leaves of the corn plant using a deep NN model and suggested an automatic classification approach.

The study by Dhivyaa et al. [32] selected the appropriate features based on bidirectional long short-term memory (Bi-LSTM) to detect plant diseases by utilizing an expanded complexity network and other dense blocks. Data from cassava disease and PlantVillage dataset have demonstrated the validity of their proposed model. According to their findings, the proposed model for the diagnosis of cassava leaf disease achieved a maximum F1 score of 95.49%.

In a study by Bhujel et al. [33], the performance of the models was improved by using a lightweight convolution neural network with several attention modules. They were able to train their models with data about tomato leaf diseases. They evaluated the models' performance (F1 score, accuracy, and recall) according to standard classification accuracy metrics. With an average accuracy of 99.69%, the convolution block focus module was the most accurate.

By using R-CNN Mask, Storey et al. [34] segmented samples to diagnose leaf and rust disease in apple orchards. For object detection, segmentation, and sickness detection, three R-CNN Mask backbones were trained and evaluated. Using the R-CNN Mask model and the ResNet-50 backbone, they were able to detect extremely small rust-covered objects.

Prabu and Chelliah [35] have developed a new method of detecting mango leaf diseases. Over 388 images of healthy and ill subjects (mango anthracnose, soot mold, and bac-

terial black spot) were selected. MobileNetV2 is used during the learning phase, and the SVM is used to classify diseases at the end.

Singh et al. [36] used the AlexNet model to identify leaf disease in maize quickly and accurately. The results were verified with PlantVillage data. *Cercospora* and Gray are two common rust-based infections covered with 99.16% accuracy in the data collection.

A lightweight convolutional neural network, RegNet, was presented by Li et al. [37] for the detection of apple leaf disease. They were able to identify five different apple leaf diseases from 2141 images of healthy and diseased apple leaves. It achieved an overall accuracy of 99.23% in the test set and 99.8% in the validation set, with a learning rate of 0.0001.

Shahoveisi et al. [38] used ML techniques to model the risk of *Sclerotinia sclerotiorum*-induced disease development in canola and dry beans. Using a broad genome correlation study [39], they previously examined the genomic regions associated with *Leptosphaeria maculans* resistance in rapeseed.

Plant leaf diseases may be consistently diagnosed using deep learning. However, TL may be used to overcome the difficulty of small datasets in plant leaf disease recognition and significantly increase the approach's accuracy. A single structure can be utilized to perform multiple tasks through ensemble learning. To solve the challenge of recognizing leaf illnesses on two different plants, the study provides research on deep transfer learning and alternate learning. Improvements were made to the residual neural network (ResNet) model, while keeping the ResNet model's convolutional layer structure. As a result, these two components of the ResNet model's classification layer can be shared across different datasets. A new fully connected layer is built to address the various classification challenges that arise while working with several data sources. The loss function and optimizer for each classifier are unique. The convolutional layer of the ResNet is fixed during formal training, allowing the TL architecture to be used for transfer learning. Formal instruction utilizes only the newly developed categories and layers that are completely interconnected. Digital cameras may be used to take on-the-spot images utilizing fruits, citrus, and rice leaves and models to diagnose leaf diseases, allowing crops to be saved before irreparable harm occurs. Although several approaches exist for diagnosing fruit leaf disease, key questions remain unresolved.

(1) Identifying and collecting critical information from fruit leaves is fairly difficult, as is distinguishing the features of particular diseases utilizing conventional image processing methods.

(2) Due to the huge variability in the features of different diseases, the attributes of disease patterns must be thoroughly studied using a diverse set of images in a smart fashion.

(3) The efficiency of ML techniques is entirely determined on the type of the hand-crafted features. As a result, feature extraction must be performed automatically in order to pick and learn the optimal collection of features for recognition.

(4) Certain deep learning models make advantage of well-established architectures such as TL structures. As a result, it employs millions of images for classification procedure. To enable immediate implementation of such models, a trade-off between computing load and accuracy must be determined.

(5) Additionally, the DL network should be trained with a large number of images to guarantee that the features are more generalized.

By employing a modified deep model and ensemble of learning based on the pool of SVM method with RBF kernel for disease detection in fruit leaves, the proposed study solved the identified research gaps. The following are some of the project's most significant contributions.

(1) A novel and distinct DL architecture has been presented in this field of research. The first goal was to use color space to improve feature classification. The second stage uses TL to learn and enhance performance by learning distinct feature maps.

(2) This is the first time a hybrid model based on the modified deep TL network and ensemble of learning has been used to diagnose disease in a significant number of fruit leaves, as far as we know.

(3) Samples from a variety of approaches were used to train the design. We trained the architecture with 90000 and validated it with 10-fold cross-validation of images.

Section 2 focuses on fruits, citrus, and rice leaf diseases, along with related datasets and processing methods. Identifying fruits, citrus, and rice leaf diseases is the subject of Section 3, which examines several concepts and methodologies. Section 4 tests the model described in this article and compares it to other models. There is a brief summary of the article in the Section 4.

2. Proposed Method

The suggested technique is depicted in Figure 1, which begins with feature extraction using a deep convolutional neural network model and ends with classification using an optimum support vector machine as the final decision layer.

2.1. Preprocessing. To accurately reproduce the appearance of colors in natural light on the image processing step, an HSV (color, saturation, and quantity) display must be used. Since the strongest color of HSV can be compared to white light, the terms "HSV" and "white light" are often used interchangeably (i.e., bright white light shining on a red object). In low light, a red object seems darker and brighter, while it appears redder and brighter in high light. We are in dire need of ensuring that no light is lost in the process; a single point source must be achieved in the light of leave images. A preprocessed RGB image is fed into the HSV converter to keep the brightness constant. The process of transforming an RGB image to an HSV image is shown visually in Figure 2.
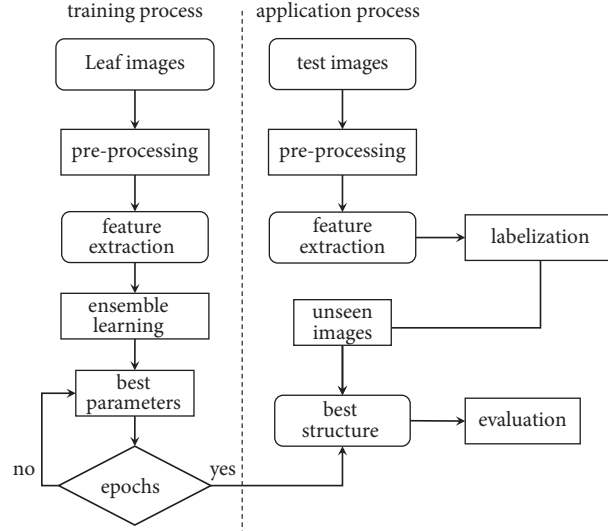
FIGURE 1: The general procedure for the suggested implementation approach is illustrated in this figure.
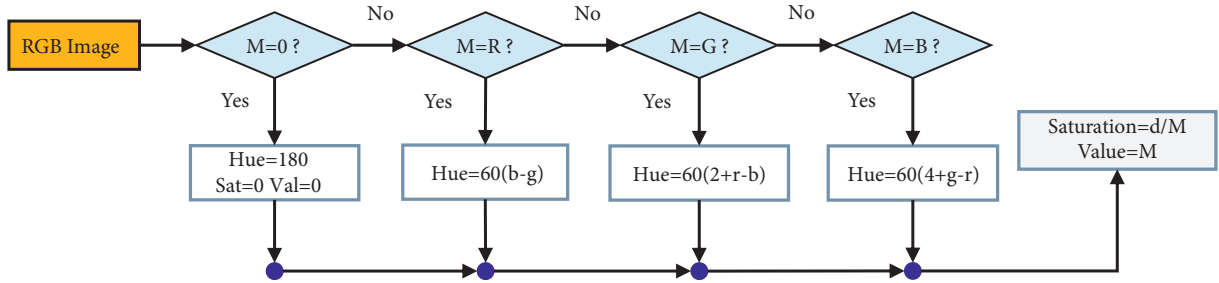
FIGURE 2: Graphic schematic from stages of RGB image to HSV image transform [40].

2.2. Convolutional Neural Network (CNN).

We use nonlinearity and a series of convolutional filters to solve equation (1). There is a level of hierarchy in a CNN. The following layers (x_j) are derived from the input signal x [41]:

$$x_j = \rho W_j x_j - 1. \tag{1}$$

However, even though W_j is a linear operator, this situation exhibits nonlinear behavior. Additionally, W_j is frequently employed in convolutions in CNNs, with ρ being either a rectifier $\max(x; 0)$ or an exponential sigmoid $[1 + \exp(-x)]^{-1}$. Convolutional filter stack W_j is assumed. As a result, the convolutions in each layer are defined as the total of the convolutions in the layer before it [41, 42].

$$x_j(u, k_j) = \rho \left(\sum_k \left(W_{j, k_j}(., k) * x_{j-1}(., k) \right)(u) \right). \tag{2}$$

In addition, a discrete convolution process known as $*$ is employed [41]:

$$(g * f)(x) = \sum_{u=-\infty}^{\infty} g(u) f(x - u). \tag{3}$$

CNN architecture optimization is a nonconvex issue. Weights W_j are frequently trained using stochastic gradient descent, with their gradients commonly determined using backpropagation. The reliance on data is a critical concern in deep learning. Deep learning models are largely reliant on massive volumes of data to be trained. In comparison to more standard machine learning algorithms, this is a significant improvement. The issue is that, in order to train the underlying data patterns, a large amount of data is required. TL can be used to address the issue of training data being distributed uniformly independent of the distribution of the test data (i.e., which motivates us to employ TL to address the problem of insufficient training data). VGG, DenseNet, and ResNet are three well-known deep learning networks that train their models using CNNs procedure.

2.3. Dilated ResNet.

ResNet's 3×3 Visual Geometry Group (VGG) full-layer design is the best on the market for transfer learning. The 3×3 output channels for each of the remaining convolutional layers are shown in Figure 3(a). After the data has been modified, it is submitted to ReLU layers. ReLU is activated following two convolutions. ResNet is based on the GoogLeNet structure as a starting point. Following the 7×7 layers are two strides and a convolution layer with up to 64, 3×3 input channels. It is feasible to connect the convolutional layer's outputs to a single input. Increasing the number of channels is one possibility. Furthermore, additional 1×1 convolution layers are also necessary. ResNet
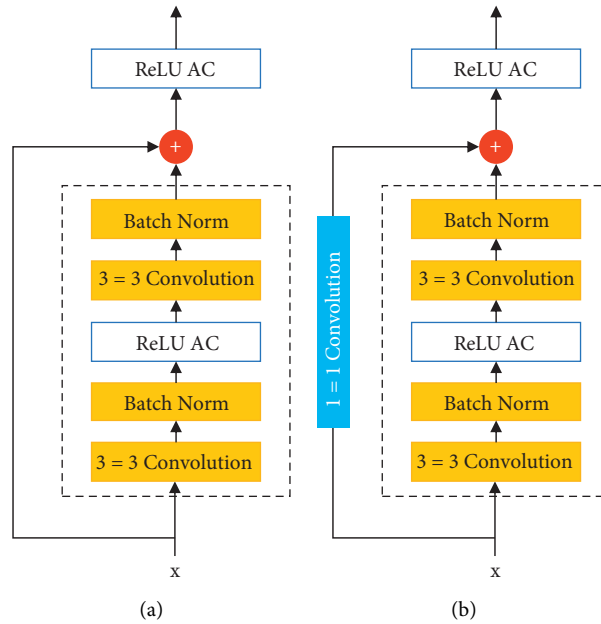
FIGURE 3: We use the ResNet residual block in the modified ResNet architecture: (a) direct mapping based on the training; (b) computation difference between the output and input.

nodes are depicted in block form in Figure 3(b). ReLU takes nonlinear input into account. This layer has no effect on the channels preceding it. Due to the small volume of diseased spots in leaf images, the dilated ResNet structure aids in the detection of images with a small volume of disease. Due to the higher-quality leaf images in the ResNet architecture, a broader range of dilated structure-based features will be available. By employing samples and preceding layers, it is feasible to preserve fine-grained features. Thus, the lightweight structure of ResNet-18 was developed by experimentation with various convolution scales and dilation rates. Concatenation hence minimizes the computational cost while increasing the quantity of fine-grained attributes. Accordingly, samples must be collected in order to combine lower-level components into higher-level maps. The bottom and top layers should be merged to improve the identification of small volumes of illness in leafs.

Each of the 64 filters illustrated in Figure 4 has the 7×7 dimension, for a total of 64. To put it simply, $Conv1$ features a significantly larger 7×7 dimension, 64 filters, and a two-step stride. The stride length of the MaxPool 2×2 is two. Following feature extraction by deep structure, we further select the feature with the lowest computational complexity. The weights of features in a neighborhood can be computed using distance measures. It is feasible to minimize the size of the required feature vectors by employing the neighborhood component analysis (NCA) [43] technique, which stands for "supervised and nonparametric." This approach allows for the enhancement and modification of the k-nearest neighbor (k-NN) structure. We choose this method to decrease the amount of feature vectors obtained from leaf images. This is because it assigns a positive weight to each feature and the feature rank can be computed using NCA.

When a desired feature reduction strategy is applied, the NCA is used to estimate the feature weights. A feature set is subdivided into overlapping blocks. As a result, it is evident that k is a collection of smaller vectors.

2.4. Ensemble Learning. The classifying leaf images with the SVM model are a significant step toward generating the ideal hyperplane as a decision surface with the highest margin for interclasses of leaf diseases. As a result of the inherent separation issues, we present the radial basis function (RBF) kernel function and its associated decision components for SVM classification. A nonlinear kernel improves the overall performance of the SVM.

$$K(x, x_i) = \exp\left(-\frac{\gamma \|x - x_i\|^2}{2\sigma^2} \right).$$ (4)

The determined RBF kernel in equation (4) contains fewer variables (C and γ), is mathematically simpler, and has fewer hyperparameters than other kernels. This is why it has gained widespread acceptance. The classification step utilizes training data to implement the SVM with RBF-based learning pool as the ensemble learning procedure. The extended ensemble learning structure is depicted in Figure 5 as a general schematic.

3. Experimental Results

This section describes the dataset and leaf image that were analyzed, as well as the model's results.

3.1. Procedure Setting. On Windows 10 PCs, the proposed solution was implemented using MATLAB R2021b. The

$$\text{Conv2} = \begin{pmatrix} 1 \times 1, 64 \\ 3 \times 3, 643 \\ 1 \times 1, 64 \end{pmatrix} \quad \text{Conv3} = \begin{pmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 64 \end{pmatrix} \quad \text{Conv4} = \begin{pmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{pmatrix} \quad \text{Conv5} = \begin{pmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{pmatrix}$$

Conv1 = 7 × 7, stride = 2, filter = 64   MaxPool = 2 × 2 , stride = 2
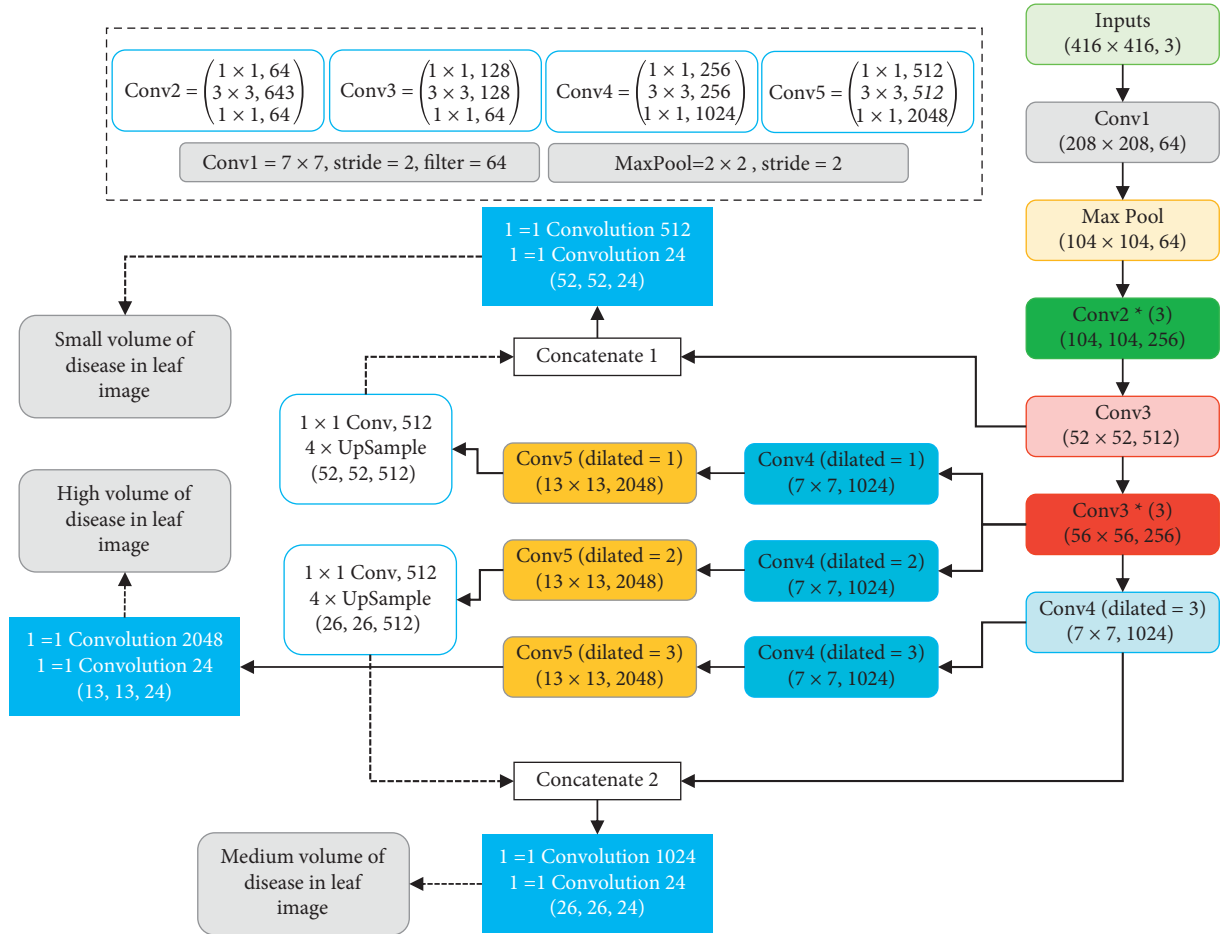
FIGURE 4: The overall architecture of modified transfer learning structure by using dilated and concatenation design.
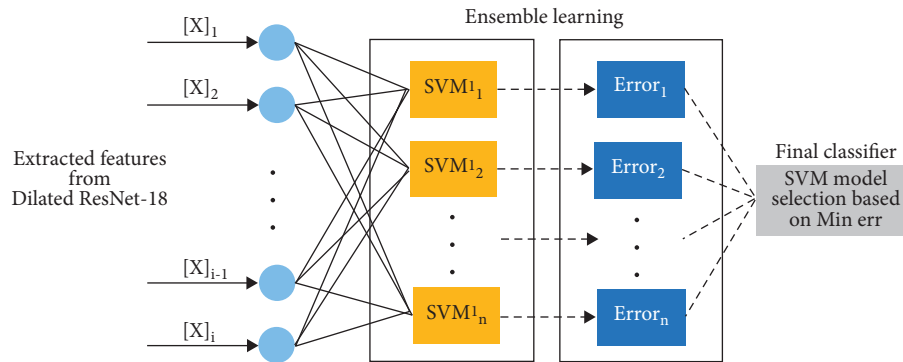
FIGURE 5: The procedure of decision-making part based on the ensemble of learning in pool of SVMs models with RBF kernel.

simulation was conducted utilizing an Intel® Core™ i5-8500 processor (single CPU), 16 GB of RAM, and 16 GB of SSD RAM on a test computer. Additionally, we used additional applications, such as SPSS. The optimal parameters of the RBF kernel are then chosen for initialization. The initial ResNet-18 model had a primary learning rate (μ) of 0.002 and an epoch range of 100–1000. To avoid overfitting, the proposed ensemble learning system determined the ideal

training iteration size by a mix of parameter tuning and early stopping.

3.2. Dataset. The images of plant leaves were created by analyzing six different types of plant leaves: apple fruit leaves, maize leaves, cotton plant leaves, grape fruit leaves, pepper leaves, and rice leaves. Figures 6 and 7 illustrate
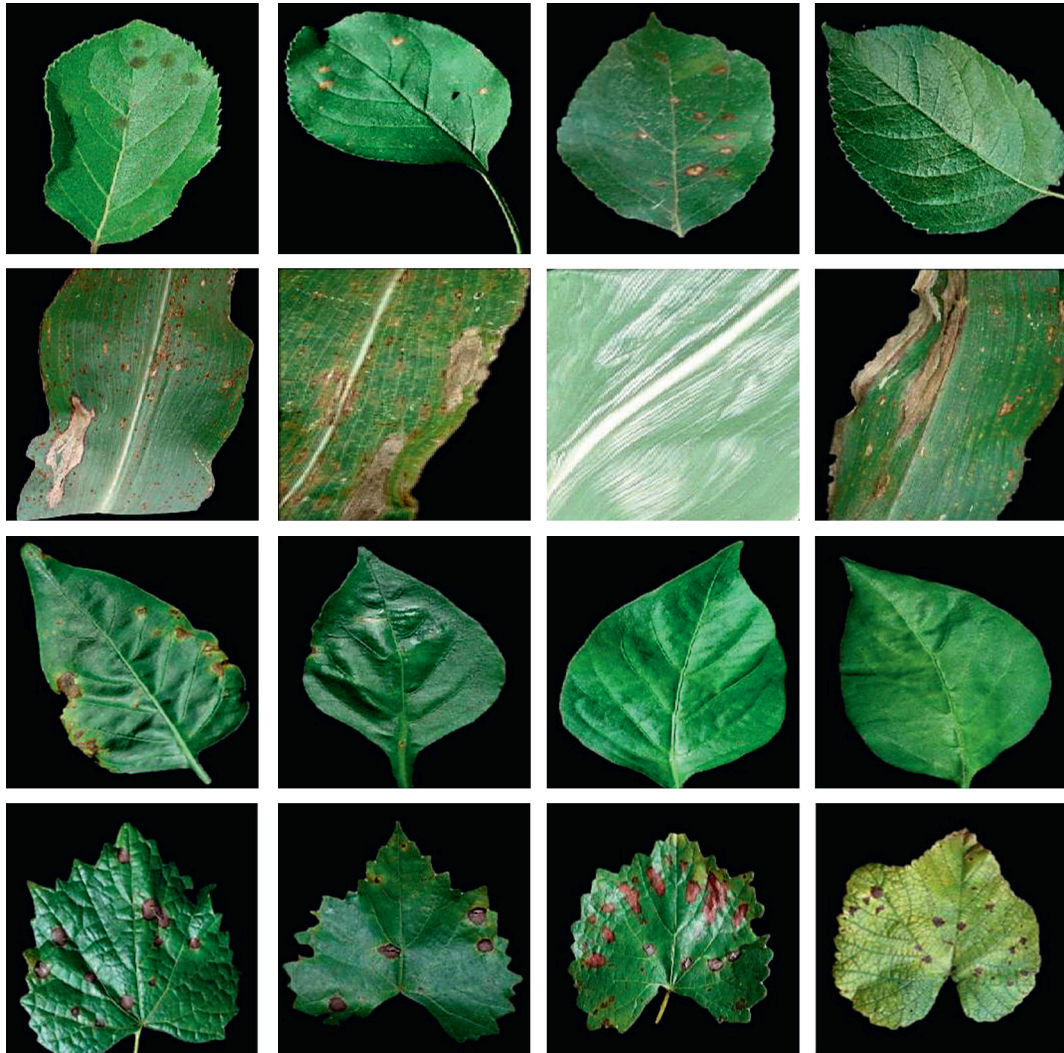
FIGURE 6: A collection of images of plant leaves, including the first row of apple fruit leaves, the second row of maize leaves, the third row of pepper leaves, and the fourth row of grape leaves, all of which have leaves with various pests. Additionally, some leaves are healthy.
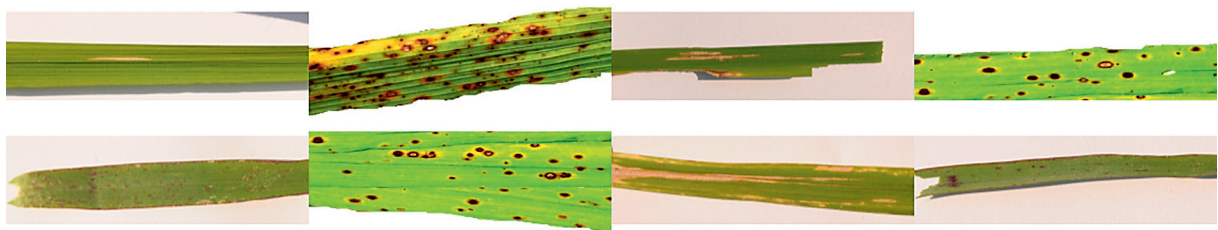
FIGURE 7: Rice leaves are divided into three different classes of bacteria, stains, and plants.

sample images, while Table 1 summarizes the number of images and other pertinent information.

3.3. Evaluations. Our early findings analyze how machine learning techniques based on ResNet structure families can be applied to improve leaf disease categorization. The feature extraction stage was carried out using an improved ResNet-18 (dResNet-18) version in ResNet structures with multiple classes and two classes, as indicated in Table 2. Additionally,

a number of previously unknown leaf diseases were identified as a result of the multiclass classification based on region of interest (ROI) part and non-ROI segment (see Table 1).

While the remainder parts of the proposed method stay unchanged, the ResNet deep transfer learning family has been in its place. Numerous parallels exist between the qualities associated with deep learning. To conduct the classification, the best SVM model is selected using an ensemble learning technique. However, the technique used

TABLE 1: This table shows the details of the leaves of the various fruits analyzed in the study.

Leaf type	Number of classes	Disease type	Dataset size	Sample		
Apple	4	Scab	3150			
		Black rot	3726			
		Cedar apple rust	1650			
		Healthy	4284			
Corn	3	Common rust	7152			
		Healthy	6972			
		Northern leaf blight	5910			
Cotton	4	*Fusarium* wilt	174			
		Myrothecium leaf spot	1062			
		Mela (soreshin)	1188			
		Areolate mildew	10482			
Grape	3	Black rot	7080			
		Black measles	8304			
		Leaf blight	6456			
Pepper	2	Bacterial spot	5982			
		Healthy	8868			
Rice	3	Bacterial leaf blight	40			
		Brown spot	40			
		Leaf smut	40			

to extract features by ResNet-101 and ResNet-164 is nearly identical.

While the proposed strategy for deep learning feature extraction outperformed other methods within the ResNet structure families in some circumstances, it is obvious that these methods occasionally produced a very small classification error.

Six different sets of gathered leaf photos are shown in Figure 8, each with its own RMSE convergence and loss function denoted by a curve and the lowest achievable value. The model's complexity is reduced by increasing the layer-wise reduction factor. As the model becomes simpler, the RMSE tends to rise. If the algorithm is trained to employ efficient features, it is feasible to improve discrimination outcomes. As a result, increasing the quantity of characteristics obtained by dilated deep transfer can help with learning. It is possible to create

useful feature maps from a leaf image with only a few iterations and a short period of time. Because of the processing time involved, we may want to avoid deeper structures for real-time or near-real-time applications (particularly during training). The confusion matrices for the integrated model's categorization across all six datasets are shown in Figure 9.

While the classification accuracy of the grape and cotton leaf datasets is lower than that of other datasets, adding the proposed extracted features significantly improves classification accuracy. In addition, 10-fold cross-validation is used to demonstrate the ability of the system to classify leaf disease. For images of healthy or sick leaves, this approach asserts a 99.1% accuracy.

Using the approach explained here, numerous low and high intensities leaf diseases can be detected in a wide variety of fruits. It is reasonable to anticipate the proper classification rate

TABLE 2: Comparison of the findings of ResNet family's error in leaf disease classification using normal and ROI images for a multiclass problem.

Leaf type	ResNet-164		ResNet-152		ResNet-101		ResNet-50		ResNet-18		dResNet-18	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Apple	0.03	0.07	0.04	0.08	0.04	0.09	0.05	0.08	0.06	0.10	**0.01**	**0.03**
Corn	0.06	0.12	0.06	0.12	0.07	0.11	0.07	0.12	0.09	0.13	**0.03**	**0.06**
Cotton	0.03	0.08	0.03	0.09	0.04	0.09	0.04	0.09	0.05	0.10	**0.02**	**0.05**
Grape	**0.04**	**0.06**	0.06	0.08	0.07	0.09	0.07	0.11	0.07	0.12	0.04	0.08
Pepper	0.03	0.06	0.04	0.07	0.04	0.07	0.04	0.08	0.05	0.08	**0.02**	**0.04**
Rice	0.005	0.01	0.01	0.015	0.008	0.02	0.01	0.02	0.01	0.03	**0.002**	**0.008**
Apple (ROI)	0.02	0.06	0.03	0.05	0.03	0.05	0.03	0.06	0.04	0.07	**0.007**	**0.02**
Corn (ROI)	0.04	0.08	0.04	0.09	0.05	0.09	0.06	0.10	0.07	0.10	**0.01**	**0.04**
Cotton (ROI)	0.01	0.06	0.02	0.07	0.02	0.08	0.03	0.08	0.03	0.09	**0.016**	**0.03**
Grape (ROI)	**0.01**	**0.05**	0.03	0.07	0.04	0.07	0.04	0.09	0.06	0.10	0.013	0.05
Pepper (ROI)	0.02	0.05	0.03	0.06	0.03	0.07	0.03	0.07	0.03	0.08	**0.009**	**0.04**
Rice (ROI)	0.004	0.008	0.005	0.018	0.006	0.02	0.008	0.02	0.009	0.02	**0.001**	**0.007**
Average	0.036	0.086	0.047	0.097	0.053	0.10	0.055	0.11	0.068	0.12	**0.009**	**0.044**

(a)

(b)

(c)

(d)

(e)

(f)
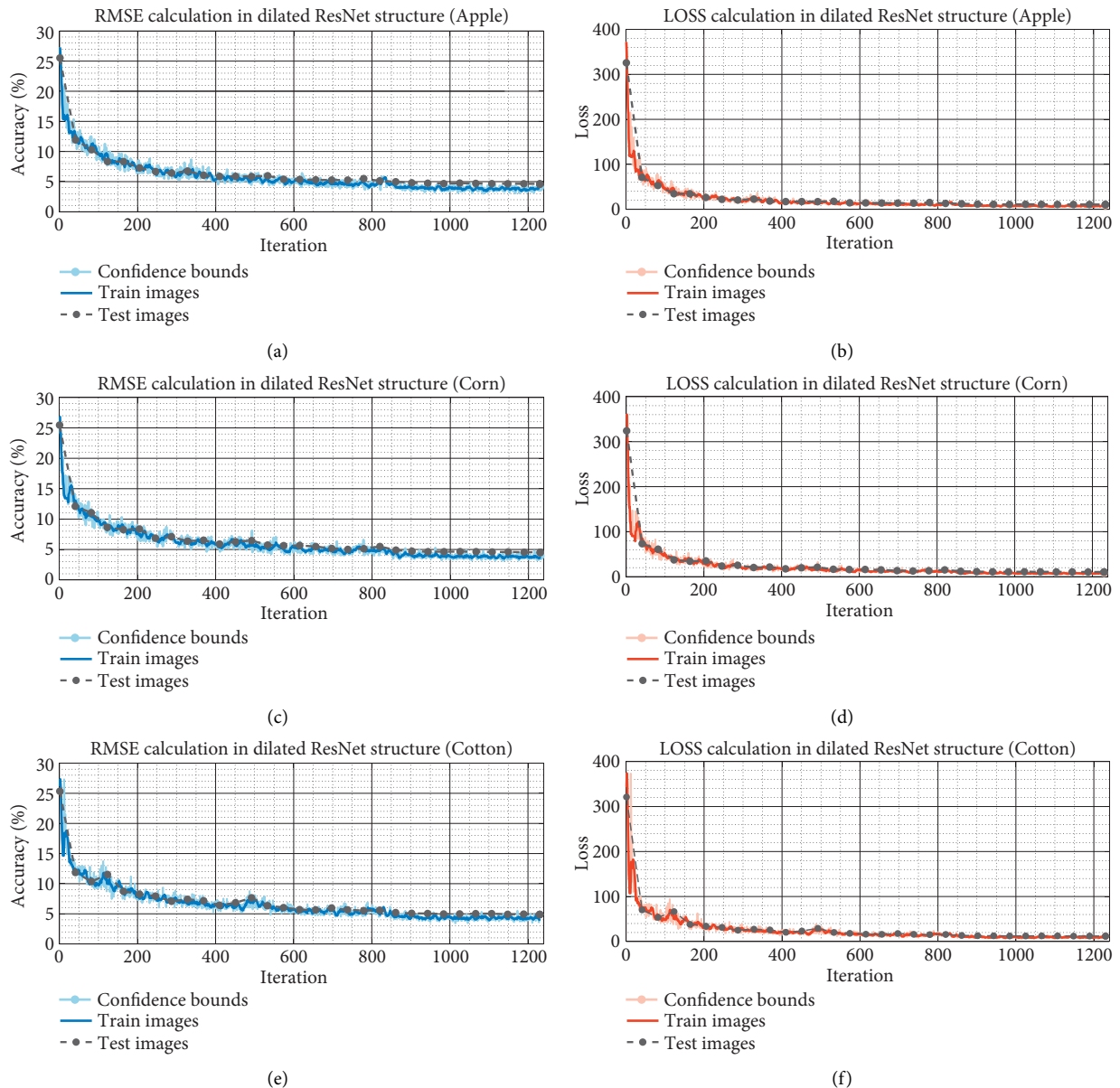
FIGURE 8: Continued.

(g)

(h)

(i)

(j)

(k)

(l)
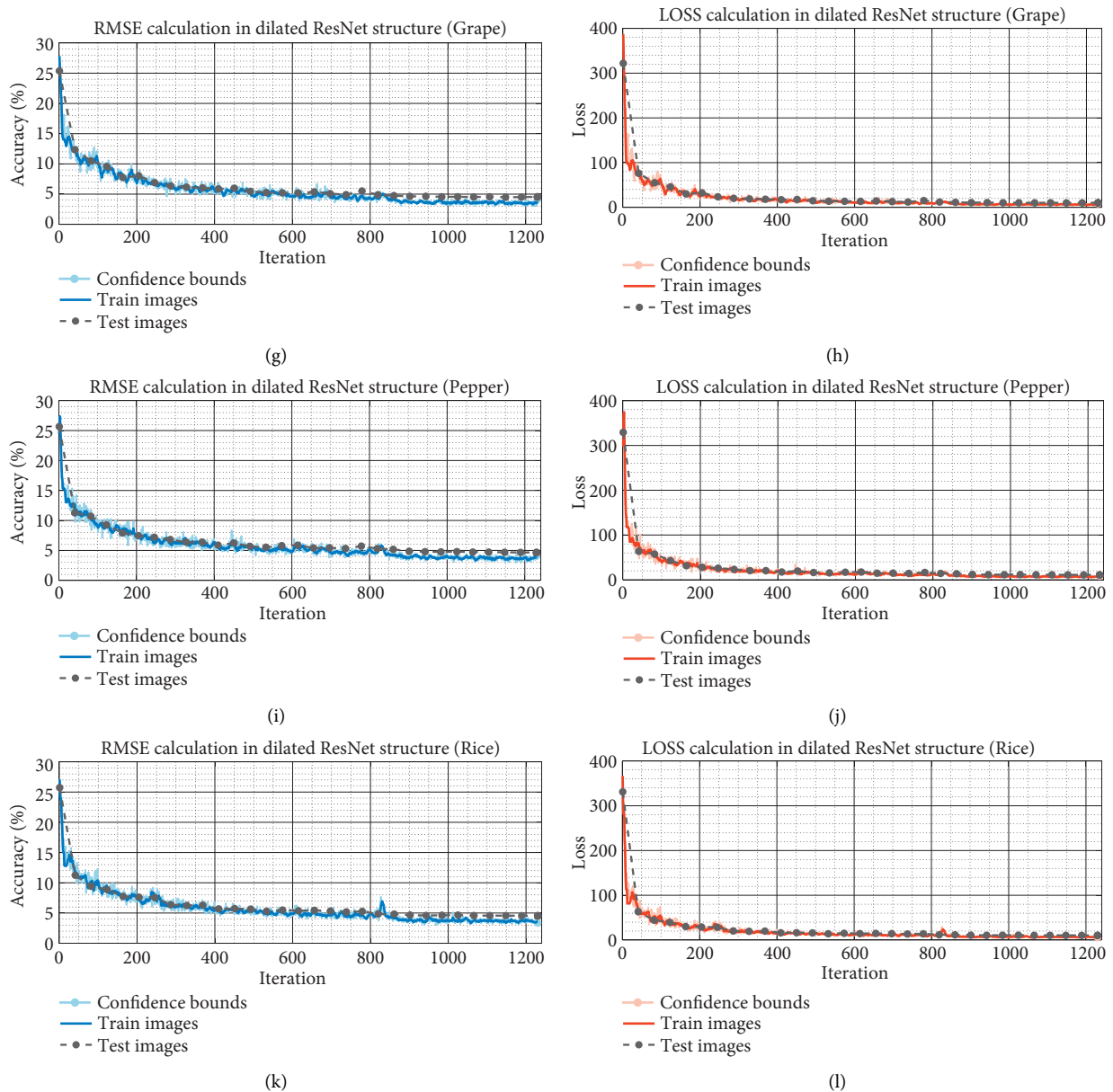
FIGURE 8: This figure depicts the RMSE convergence and loss function of the feature extraction section for six sets of acquired leaf images. (a, c, e, g, i, k) are the RMSE calculation and (b, d, f, h, j, l) are the loss computation for apple, corn, cotton, grape, pepper, and rice leaf images, respectively.

because it consistently obtains a 99.1% accuracy rate in trials involving two classes in a range of situations. Our model achieves a precision score of 100% on each rice and apple leaves image. The proposed deep model and effective learning approaches for feature extraction may have resulted in appropriate sensitivity. The results show that the proposed method for diagnosing and classifying leaf diseases tends to be reliable. When a leaf image is complicated, as seen in Figure 9, various techniques' accuracy and precision rates often decrease (e.g., especially when the images are noisy or non-RoI image is selected as based for processing). The approach for analyzing leaf images, on the other hand, is consistent and reliable, making it a cost-effective method.

4. Discussion

This study employed a technique known as model engineering (ME) learning. There are various approaches to train a model for plant disease recognition; however, our model was trained using the dilated ResNet-18 framework for deep transfer learning.

Our learning algorithms were all constructed using identical data and loss functions. The ME model described in this article was applied to a range of fruit leaf images. The RMSE and loss error change curves for each model's training set are displayed in Figure 8. Figure 10 illustrates the accuracy change for various TL model
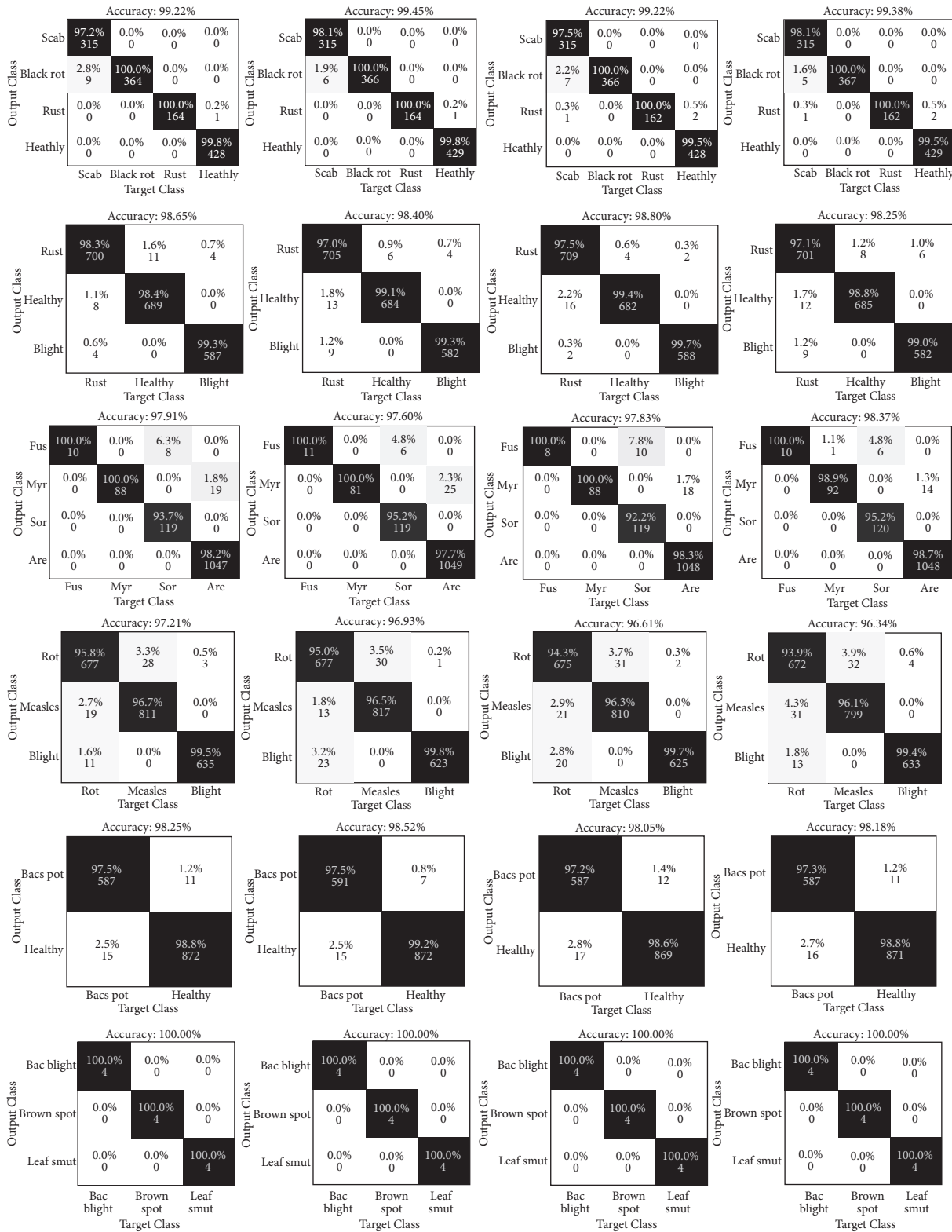
FIGURE 9: Confusion matrix: the suggested approach is used to diagnose leaf disease in a variety of hitherto unexplored datasets. The method's output is displayed four times at random. The first row represents apples, the second row represents corn, the third row represents flax, the fourth row represents grapes, the fifth row represents peppers, and the sixth row represents rice, accordingly.
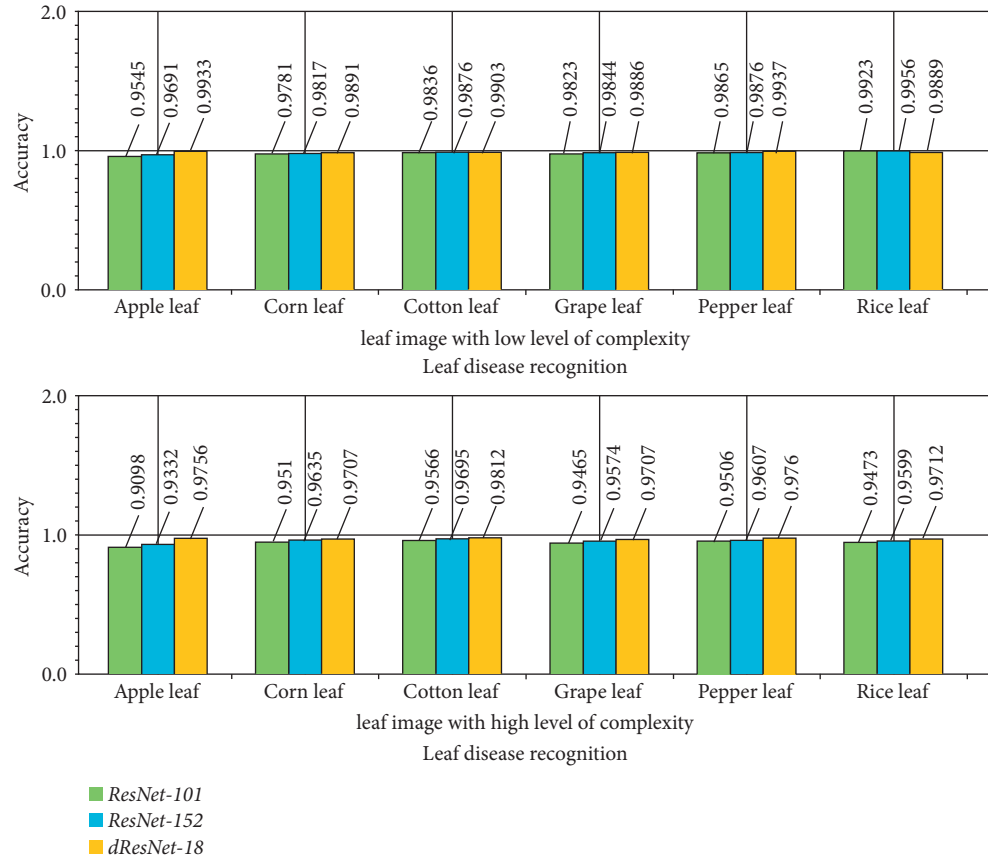
FIGURE 10: This figure illustrates the accuracy change for various TL model verification sets based on the low and high levels of complexity.

verification sets based on the low and high levels of complexity.

The accuracy rates for the different decision-making part based on various test sets are shown in Table 3 based on k-nearest neighbor (k-NN), decision tree (DT), feedforward neural network (NN), SVM with linear kernel (SVM-L), SVM with RBF kernel (SVM-RBF), and Ensemble SVM-RBF (E-SVM-RBF).

The dilated learning model outperforms the typical ResNet-18 design in terms of stability. This model has an average accuracy of 98.5% on the test set for leaf disease recognition models. A test set accuracy of 97.93% is less than the suggested structure's accuracy of 97.93% in recognizing Grape or Cotton leaf diseases. For instance, our new deep learning model outperforms existing deep learning models by an average of 1 to 2 percentages. The addition of acceptable noise while learning a task aids in the improvement of models, which is analogous to removing unnecessary steps from the process. In contrast to conventional learning, interaction can assist in preventing the overfitting of different leaf image conditions from collapsing. As a result, the proposed ME learning technique can aid in the achievement of numerous, relatively related objectives. For all leaf disease recognition, the proper epochs were used to optimize the training time in the procedure. This resulted in a more rapid and constant drop in the recognition model's accuracy. Our

TABLE 3: The comparison of the different decision-making part based on various test sets.

Leaf image type	k-NN (%)	DT (%)	NN (%)	SVM-L (%)	SVM-RBF (%)	E-SVM-RBF (%)
Apple	97.56	97.31	96.11	95.48	98.13	99.47
Corn	97.08	96.83	95.78	94.11	97.29	98.69
Cotton	95.71	96.10	95.39	96.73	97.76	98.91
Grape	96.18	96.51	96.26	95.31	97.07	98.74
Pepper	97.09	96.65	95.62	96.13	97.87	99.12
Rice	98.31	98.87	97.91	96.44	98.76	100

ME model surpasses existing feature extraction-based deep TL models.

There are multiple distinct methods for incorporating pretrained structures into TL models. Minor changes are made using the TL approach. This was no longer essential as a result of the ensemble classification approach. Model reuse is distinguished by the fact that it uses the model's structure but not its pretrained parameters. We processed images of leaves on a variety of fruits using the approach indicated above. According to the results in Table 4, when using the DL structure with a limited number of epochs in the training stage for creating proper features with the appropriate metrics computation, the hybrid model is capable of recognizing leaf disease with a maximum accuracy of 99.1%.

TABLE 4: Comparison of proposed method with other deep and classical learning structures.

Reference	Type of leaf	Model	Accuracy (%)	Computational complexity
Chen et al. [10]	Rice	VGGNet	92.00	High
Sharma et al. [25]	Tomato	S-CNN and F-CNN	98.30	High
Atila et al. [44]	Plant leaf	EfficientNet	96.18	High
Kaur et al. [45]	Grape	Hy-CNN	98.70	Medium
Ji et al. [46]	Grape	United model	98.20	High
Gadekallu et al. [47]	Plant leaf	Whale and DL	95.10	Medium
Azimi et al. [48]	Crop	FCNN and SCNN	92.01	High
Joshi et al. [49]	Coffee	Deep CNN	98.00	High
Paymode and Malode [50]	Grape and tomato	VGG	98.40 and 95.71	High
Jiang et al. [51]	Rice and wheat	VGG16, DenseNet-121, and ResNet-50	97.22 and 98.75	High
Nandhini and Ashokkumar [52]	Plant leaf	DenseNet-121 and optimization algorithm	98.70	High
Our model	Plant leaf	Dilated TL and ensemble learning	99.10	Medium

Our model, which was taught to recognize leaves in a variety of fruits, demonstrated a reliable effect that might be used as actual smart technology in agriculture. Other studies looked into whether DL-based models are unique in their ability to distinguish between diseases from different time periods. The table below compares the dResNet-18 and other decision-making structures to a variety of feature extraction and ensemble learning models. To detect, recognize, and describe leaf disease in plants, many deep learning models are used. Various CNN models, such as S-CNN and F-CNN [17], EfficientNet [43], Hy-CNN [45], and the united deep learning model [46], have been applied to various plant leaf images by a number of authors. On plant leaves, a hybrid analytic model [47] obtains 95.1 percent precision, while other models achieve 92.01% precision [48]. The classification of images of coffee leaves is improved by texture image analysis [49]. In addition, our approach is a fast method that has less computational complexity than other similar methods in leaf disease classification.

As a result of its many advantages, deep learning in data science can lead to more efficient processing models [53]. A decision-making approach proposed in this study could be of significant help in classifying leaf diseases. By using deep learning, accuracy can be continuously increased and knowledge can be continuously gained without supervision. Additionally, data analysts will benefit from more accurate and concise results as a result of this initiative.

5. Conclusion

This article analyzes and provides a dataset of six different types of crop disease leaves. The deep CNN-based dilated ResNet-18 model is trained and evaluated using data processing, preprocessing of the dataset, training, and experimental procedure. The proposed model is constructed and tested to determine whether it can improve on the outcomes obtained thus far. Various datasets indicate that the metrics parameters are stronger and higher than those of other techniques. As a result, our proposed research work increased the accuracy of various leaf images of various fruits by 99.1 percent, including Apple, Corn, Cotton, Grape, Pepper, and Rice. We were able to achieve the highest level of performance possible with our model when it came to on-

field crops, leaf images, and disease classification and analysis. Agriculture and food production are the core objectives of university and government-funded research. It is intended that real datasets comprising a variety of leaf crops would be collected and processed for usage in deep learning models in the future. Crop images will significantly benefit from the future usage of hybrid models based on improved TL architecture and traditional models. Farmers gain from our efforts because we encourage and stimulate agriculture, which boosts agricultural revenue and contributes to the development of strong countries. Additionally, in the future, the authors of this study will attempt to reduce the algorithm's computational complexity and execution time.

Data Availability

All the data and codes are available from corresponding authors.

Ethical Approval

The research meets all applicable standards with regard to the ethics of experimentation and research integrity.

Conflicts of Interest

The author has no conflicts of interest to declare.

References

[1] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017.

[2] V. K. Vishnoi, K. Kumar, and B. Kumar, "Plant disease detection using computational intelligence and image processing," *Journal of Plant Diseases and Protection*, vol. 128, no. 1, pp. 19–53, 2021.

[3] J. Xiong, D. Yu, S. Liu, L. Shu, X. Wang, and Z. Liu, "A review of plant phenotypic image recognition technology based on deep learning," *Electronics*, vol. 10, no. 1, p. 81, 2021.

[4] I. Perez-Borrero, D. Marin-Santos, M. J. Vasallo-Vazquez, and M. E. Gegundez-Arias, "A new deep-learning strawberry instance segmentation methodology based on a fully

convolutional neural network," *Neural Computing & Applications*, vol. 33, pp. 15059–15071, 2021.

[5] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, "A review of convolutional neural network applied to fruit image processing," *Applied Sciences*, vol. 10, no. 10, Article ID 3443, 2020.

[6] M. Azadbakht, D. Ashourloo, H. Aghighi, S. Radiom, and A. Alimohammadi, "Wheat leaf rust detection at canopy scale under different LAI levels using machine learning techniques," *Computers and Electronics in Agriculture*, vol. 156, pp. 119–128, 2019.

[7] M. A. Ebrahimi, M. H. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vision-based pest detection based on SVM classification method," *Computers and Electronics in Agriculture*, vol. 137, pp. 52–58, 2017.

[8] J. García, C. Pope, and F. Altimiras, "A distributed-means segmentation algorithm applied to lobesia botrana recognition," *Complexity*, vol. 2017, Article ID 5137317, 14 pages, 2017.

[9] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.

[10] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Computers and Electronics in Agriculture*, vol. 173, Article ID 105393, 2020.

[11] W. Xie, G. Ma, F. Zhao, H. Liu, and L. Zhang, "PolSAR image classification via a novel semi-supervised recurrent complex-valued convolution neural network," *Neurocomputing*, vol. 388, pp. 255–268, 2020.

[12] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: a survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

[13] Z. Li, R. Guo, M. Li, Y. Chen, and G. Li, "A review of computer vision technologies for plant phenotyping," *Computers and Electronics in Agriculture*, vol. 176, Article ID 105672, 2020.

[14] A. I. Mourikis, R. Kalamatianos, I. Karydis, and M. Avlonitis, "A survey on the use of the internet of multimedia things for precision agriculture and the agrifood sector," *Engineering Proceedings*, vol. 9, no. 1, p. 32, 2021.

[15] P. Kartikeyan and G. Shrivastava, "Review on emerging trends in detection of plant diseases using image processing with machine learning," *International Journal of Computer Application*, vol. 975, p. 8887, 2021.

[16] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, "Detection of rice plant diseases based on deep transfer learning," *Journal of the Science of Food and Agriculture*, vol. 100, no. 7, pp. 3246–3256, 2020.

[17] F. Jiang, Y. Lu, Y. Chen, D. Cai, and G. Li, "Image recognition of four rice leaf diseases based on deep learning and support vector machine," *Computers and Electronics in Agriculture*, vol. 179, Article ID 105824, 2020.

[18] J. Shin, Y. K. Chang, B. Heung, T. Nguyen-Quang, G. W. Price, and A. Al-Mallahi, "A deep learning approach for RGB image-based powdery mildew disease detection on strawberry leaves," *Computers and Electronics in Agriculture*, vol. 183, Article ID 106042, 2021.

[19] Z. Gao, L. R. Khot, R. A. Naidu, and Q. Zhang, "Early detection of grapevine leafroll disease in a red-berried wine grape cultivar using hyperspectral imaging," *Computers and Electronics in Agriculture*, vol. 179, Article ID 105807, 2020.

[20] M. Long, C. Ouyang, H. Liu, and Q. Fu, "Image recognition of Camellia oleifera diseases based on convolutional neural network & transfer learning," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 34, no. 18, pp. 194–201, 2018.

[21] X. Nie, L. Wang, H. Ding, and M. Xu, "Strawberry verticillium wilt detection network based on multi-task learning and attention," *IEEE Access*, vol. 7, pp. 170003–170011, 2019.

[22] A. Dobrescu, M. V. Giuffrida, and S. A. Tsaftaris, "Doing more with less: a multitask deep learning approach in plant phenotyping," *Frontiers of Plant Science*, vol. 11, p. 141, 2020.

[23] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Solving current limitations of deep learning based approaches for plant disease detection," *Symmetry*, vol. 11, no. 7, p. 939, 2019.

[24] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," *Applied Soft Computing*, vol. 86, Article ID 105933, 2020.

[25] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Information Processing in Agriculture*, vol. 7, no. 4, pp. 566–574, 2020.

[26] G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 27–34, 2021.

[27] J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, "A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network," *Computers and Electronics in Agriculture*, vol. 154, pp. 18–24, 2018.

[28] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.

[29] J. G. Arnal Barbedo, "Plant disease identification from individual lesions and spots using deep learning," *Biosystems Engineering*, vol. 180, pp. 96–107, 2019.

[30] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Computers and Electronics in Agriculture*, vol. 158, pp. 20–29, 2019.

[31] A. K. Singh, B. Ganapathysubramanian, S. Sarkar, and A. Singh, "Deep learning for plant stress phenotyping: trends and future perspectives," *Trends in Plant Science*, vol. 23, no. 10, pp. 883–898, 2018.

[32] C. R Dhivyaa, N. Kandasamy, and S. Rajendran, "Integration of dilated convolution with residual dense block network and multi-level feature detection network for cassava plant leaf disease identification," *Concurrency and Computation: Practice and Experience*, Article ID e6879, 2022.

[33] A. Bhujel, N.-E. Kim, E. Arulmozhi, J. K. Basak, and H.-T. Kim, "A lightweight Attention-based convolutional neural networks for tomato leaf disease classification," *Agriculture*, vol. 12, no. 2, p. 228, 2022.

[34] G. Storey, Q. Meng, and B. Li, "Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture," *Sustainability*, vol. 14, no. 3, Article ID 1458, 2022.

[35] M. Prabu and B. J. Chelliah, "Mango leaf disease identification and classification using a CNN architecture optimized by crossover-based levy flight distribution algorithm," *Neural Computing & Applications*, pp. 1–14, 2022.

[36] R. K. Singh, A. Tiwari, and R. K. Gupta, "Deep transfer modeling for classification of maize plant leaf disease," *Multimedia Tools and Applications*, pp. 1–17, Springer, Berlin, Germany, 2022.

[37] L. Li, S. Zhang, and B. Wang, "Apple leaf disease identification with a small and imbalanced dataset based on lightweight convolutional networks," *Sensors*, vol. 22, no. 1, p. 173, 2022.

[38] F. Shahoveisi, M. Riahi Manesh, and L. E. Del Río Mendoza, "Modeling risk of Sclerotinia sclerotiorum-induced disease development on canola and dry bean using machine learning algorithms," *Scientific Reports*, vol. 12, no. 1, pp. 864–910, 2022.

[39] S. Mansouripour, A. Oladzad, F. Shahoveisi et al., "Identification of genomic regions associated with resistance to blackleg (Leptosphaeria maculans) in canola using genome wide association study," *European Journal of Plant Pathology*, vol. 161, no. 3, pp. 693–707, 2021.

[40] K. Rezaee, J. Haddadnia, and A. Delbari, "Modeling abnormal walking of the elderly to predict risk of the falls using Kalman filter and motion estimation approach," *Computers & Electrical Engineering*, vol. 46, pp. 471–486, 2015.

[41] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, Article ID 1995, 1995.

[42] M. S. Anari, K. Rezaee, and A. Ahmadi, "TraitLWNet: a novel predictor of personality trait by analyzing Persian handwriting based on lightweight deep convolutional neural network," *Multimedia Tools and Applications*, pp. 1–21, 2022.

[43] J. Kang, R. Fernandez-Beltran, S. Liu, and A. Plaza, "Toward tightness of scalable neighborhood component analysis for remote-sensing image characterization," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[44] Ü Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecological Informatics*, vol. 61, Article ID 101182, 2021.

[45] P. Kaur, S. Harnal, R. Tiwari et al., "Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction," *Sensors*, vol. 22, no. 2, p. 575, 2022.

[46] M. Ji, L. Zhang, and Q. Wu, "Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 418–426, 2020.

[47] T. R. Gadekallu, D. S. Rajput, M. P. K. Reddy, K. Lakshmanna, S. Bhattacharya, and S. Singh, "A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU," *Journal of Real-Time Image Processing*, vol. 18, no. 4, pp. 1383–1396, 2021.

[48] S. Azimi, T. Kaur, and T. K. Gandhi, "A deep learning approach to measure stress level in plants due to nitrogen deficiency," *Measurement*, vol. 173, Article ID 108650, 2021.

[49] R. C. Joshi, M. Kaushik, M. K. Dutta, A. Srivastava, and N. Choudhary, "VirLeafNet: automatic analysis and viral disease diagnosis using deep-learning in Vigna mungo plant," *Ecological Informatics*, vol. 61, Article ID 101197, 2021.

[50] A. S. Paymode and V. B. Malode, "Transfer learning for multicrop leaf disease image classification using convolutional neural networks VGG," *Artificial Intelligence in Agriculture*, vol. 6, 2022.

[51] Z. Jiang, Z. Dong, W. Jiang, and Y. Yang, "Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning," *Computers and Electronics in Agriculture*, vol. 186, Article ID 106184, 2021.

[52] S. Nandhini and K. Ashokkumar, "An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm," *Neural Computing & Applications*, pp. 1–22, 2022.

[53] K. Rezaee, S. Savarkar, X. Yu, and J. Zhang, "A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals," *Biomedical Signal Processing and Control*, vol. 71, Article ID 103161, 2022.

Hindawi

Research Article

Intelligent Manufacturing of New Energy Vehicles and Financial Market Hedging Based on Soft Computing

Peizhi Zhao [ID]

Department of Accounting, Southampton Business School, University of Southampton, Southampton SO17 1BJ, UK

Correspondence should be addressed to Peizhi Zhao; pz1u20@soton.ac.uk

Globally, the energy supply in the market is tight and the oil price fluctuates sharply. With the increasing degree of environmental pollution, both developed and developing countries pay special attention to the development of new energy, and energy conservation and emission reduction have been put on the agenda. As new energy vehicles have the advantages of energy conservation and environmental protection, they are strongly supported by governments. Many countries regard new energy vehicles as substitutes for traditional vehicles, and their development has ushered in a good opportunity. Under the background that the government and enterprises pay more attention to the development of new energy vehicles, it is of positive significance to study the intelligent manufacturing enterprises of new energy vehicles in a province. This paper studies the manufacturing of new energy vehicles through soft computing. Next, this paper also analyzes the financial market. Enterprise risk is closely related to investor income and social healthy development, which is a hot issue concerned by the public and the government. However, in recent years, there have been a series of cases in which enterprises have suffered huge losses using "hedging" to avoid operational risks, which has aroused public concern about the motivation of enterprises to use derivative financial products. Based on the leverage characteristics of derivative financial instruments, scholars believe that the high leverage profit in the capital market is the main incentive for enterprises to use hedging for hedging speculation. Existing studies pay more attention to the consequences of enterprise hedging, and they analyze the impact of enterprise hedging, focus on the result level after enterprise hedging, and seek the significance of enterprise hedging. Through the research on soft computing, this paper will apply it to the intelligent manufacturing of new energy vehicles and promote the development of hedging in China's financial market

1. Introduction

After decades of development, the traditional automobile industry has become more technologically mature and has made great contributions to facilitating people's travel. However, with the development of the economy and the popularization of automobiles, the following two problems have also appeared: on the one hand, oil is the "blood" of the development of traditional automobiles, and it is a nonrenewable resource [1, 2]. If oil is used as the power source of automobiles for a long time, the development of cars becomes unsustainable, and the problem of fuel shortage is highly valued by governments of various countries [3]. On the other hand, traditional cars produce a lot of waste during use, causing serious damage to the environment, and car

exhaust pollution has also been criticized by people [4, 5]. Environmentalists all over the world are calling for environmental protection and green and low-carbon travel [6]. Therefore, countries all over the world are paying more attention to energy conservation and emission reduction [7]. Combining the above two factors, to solve the problem of environmental pollution, new energy vehicles based on soft computing will be an important way to solve this problem, and it has also become an important direction for the green development and sustainable development of new energy vehicle intelligent manufacturing [8]. Next, this article analyzes the financial market. In the context of the integrated development of the global financial system, financial liberalization and economic globalization are becoming more important. China's financial market and the international

financial instrument market are becoming closer [9]. Therefore, for Chinese companies, it is difficult to avoid risks caused by the price fluctuations of raw materials, products, and the financial environment [10, 11]. Companies are affected by the large fluctuations in the prices of bulk goods and upstream raw materials, and the company's profits have become more uncertain, causing huge losses to the company, and even endangering the company's operations [12]. Affected by the financial crisis and economic globalization, the price fluctuations in the bulk commodity market are not a short-term phenomenon [13]. How to deal with the impact of price fluctuations is the key to a company's sustainable operation [14]. For stable operations, companies should constantly adjust their business models to ensure that the importance of accurately identifying, grasping, and effectively avoiding risks brought by price fluctuations, and making effective use of the futures market for companies risks are hedged [15]. However, the huge losses caused by the use of hedging by enterprises are caused by the use of derivative financial instruments [16]. Based on the important management foundation of the enterprise, the perspectives of senior management and internal governance and the important external management foundation of the company are discussed from the macroperspective of the use of hedging [17]. Motivation issues help companies make decisions on derivative financial products, such as hedging, and avoid corporate risks [18].

2. Related Work

The literature shows the importance of studying the financial performance evaluation of new energy vehicles in a certain province and strives to provide suggestions for the development of the new energy vehicle industry in a province. Part of the research puts forward the improvement suggestions of the quality management team; the visual interface and system functions of the complete SPC software are designed. At the same time, it provides a software basis for the realization of the previously designed quality control system, also determines the thermal resistance of key parameters according to customer feedback and production performance, and determines the key process of heat sink dispensing process and DBC (ceramic copper sheet) cutting and mounting process. Through the 5M1E fishbone diagram analysis method, the cause of the defect was deeply analyzed, and the improvement plan, the adverse reaction plan, and the three-stage plan of SPC real-time monitoring were established. The SPC monitoring of thermal resistance parameters, glue quantity and DBC (ceramic copper sheet) placement pressure parameters, and other parameters were gradually realized, and the SPC control chart of thermal resistance parameters was drawn. There are also some literatures that propose the construction of a financial performance indicator system, which lays the foundation for the following AHP. The literature proposes to build a financial performance indicator system, which lays the foundation for the following level analysis. The literature proposes the construction of a financial performance indicator system, which lays the foundation for the analysis of the hierarchy below. At the same time, it pays attention to the benefits and problems of the company's

hedging scheme, as well as the matters that should be paid attention to in hedging. And the Shanghai copper option is introduced, by designing two optimization schemes for hedging; one is using the Shanghai copper option alone and the other by combining the Shanghai copper future and the Shanghai copper option, and forecasting the scheme. After analyzing the advantages and existing problems of the G enterprise hedging scheme, it puts forward some matters that G enterprise should pay attention to in the future hedging. The literature proposes a method of matching car styling features according to user needs and builds a multilevel labeling car styling library. The literature shows case studies on user demand feature extraction, car styling feature matching, and car styling intelligent generation.

3. Soft Computing and Intelligent Manufacturing Technology of New Energy Vehicles

3.1. Soft Computing.
Soft computing is a technique that increases actual uncertainty on the basis of hard computing to obtain an approximate solution or optimal solution to a practical problem to achieve the best solution. Soft computing technology can be applied to the research of product image design, and it can convert physical phenomena in a certain process or level of complex problems into mathematical models. At present, domestic and foreign scholars' research on soft computing methods mainly focuses on artificial intelligence and predictive models. In terms of product design, especially product images, there are few pieces of research on soft computing. The research on product image design based on computing methods in this article has certain theoretical and practical value.

As shown in Figure 1, soft computing is a set of methods used to infer and learn human thoughts, process fitting models, and provide solutions to complex problems. It allows uncertainty and incomplete actual values to exist, and this method can better solve the uncertainty problem that is common in reality.

3.2. Intelligent Manufacturing of New Energy Vehicles

3.2.1. Algorithm Model.
The TFIDF algorithm is one of the classic algorithms of keyword extraction technology. The central idea is to extract keywords in the text according to the statistical information of word frequency and calculate the weight of the corresponding keywords. The high word frequency information of a certain keyword in the text and the low word frequency information of other texts reflect the keyword's ability to distinguish other texts. Therefore, keywords can be used as keywords in the text. The specific definition of TFIDF is in equations (1) and (2).
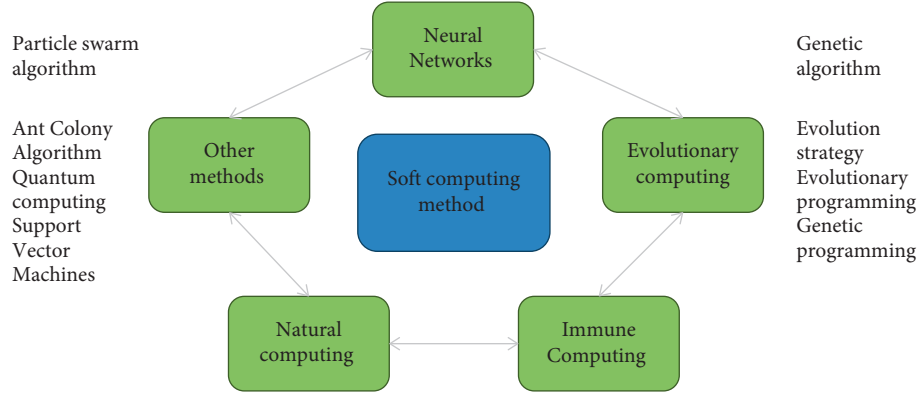
FIGURE 1: Soft computing method.

$$\begin{cases} \text{TF}_{ij} = \dfrac{n_{ij}}{\sum_k n_{kj}}, \\[4mm] \text{IDF}_i = \log \dfrac{|D|}{|D_i| + 1}, \end{cases} \tag{1}$$

$$\begin{cases} w_{ij} = \text{TF}_{ij} \times \text{IDF}_{ij}, \\[4mm] w_{ij} = \dfrac{w_{ij}}{\sqrt{\sum_k w_{kj}}}. \end{cases} \tag{2}$$

The principle of the cosine similarity calculation method is to use the cosine value of the angle between the two vectors in the vector space as the difference between individual objects. The specific definition is given in equations (3) and (4).

$$\begin{cases} A = \{a_1, a_2, a_3, \ldots, a_n\}, \\ B = \{b_1, b_2, b_3, \ldots, b_m\}. \end{cases} \tag{3}$$

From equation (3), we can see that A and B represent two independent objects, and each object contains N-dimensional features.

$$\text{sim}(A, B) = \cos\theta \frac{\vec{a} \cdot \vec{b}}{\|a| \cdot |b\|}. \tag{4}$$

It can be seen from formula (4) that the method of calculating cosine similarity is to regard the object as a coordinate point in the N-dimensional coordinate system.

To continuously optimize the generated result image, a total loss function is defined, which is used to perform gradient descent to obtain a result image that matches the original image features. The formula of the total loss function is as follows:

$$L_{\text{total}}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{\text{content}}(\vec{p}, \vec{x}) + \beta L_{\text{style}}(\vec{a}, \vec{x}). \tag{5}$$

Among them, L_{content} is the content loss function, and L_{style} is the style loss function. The specific definitions are given in equations (6) and (7).

$$L_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2}\sum\left(F_{ij}^l - P_{ij}^l\right)^2. \tag{6}$$

Use the original image to calculate the convolution feature P_{ij}^l and the convolution of the l layer.

$$L_{\text{style}}(\vec{p}, \vec{x}, l) = \frac{1}{4N_l^2 M_l^2}\sum\left(A_{ij}^l - G_{ij}^l\right)^2. \tag{7}$$

3.2.2. Data Processing. Word frequency information is a measure of the frequency of words appearing in a document, and the specific definition is as follows:

$$p(s) = p(w_1, w_2, \ldots, w_n) = \prod_{i=1}^{n} p(w_i|w_1, w_2, \ldots, w_{i-1}). \tag{8}$$

Mutual information statistics originally originated from the information theory, reflecting the connection between words. In the field of natural language processing, mutual information usually represents the internal cohesion between words, reflecting whether words can form new words. It is used to identify unregistered words and eliminate ambiguous words, and the specific definition of mutual information is as follows:

$$M(A, B) = \log_2 \frac{P(A, B)}{P(A)P(B)}. \tag{9}$$

However, upon considering the cohesion of the words alone, it is easy to split the words. The "keywords" in the "keywords" have a high mutual information value in the document, however, they are not unregistered words. Therefore, we need to introduce the concept of boundary freedom. The core idea of boundary freedom is to judge whether words can be used freely in each part of a sentence. Here, the minimum value of the number of words appearing on the left side of the phrase and the number of words appearing on the right side of the phrase is set as value of the boundary degrees of freedom. Equation (10) is defined as follows:

$$F(L, R) = F_{\text{minlen}}\left(L(w_1, w_2 \ldots w_n), R(w_1, w_2 \ldots w_m)\right). \tag{10}$$

The word span reflects the description range of the descriptive vocabulary. The more the span, the more important the word and the stronger the whole. The calculation method is as follows:

$$W_{\text{span}} = \frac{|l_i|}{|L|}. \tag{11}$$

Following the above-mentioned vocabulary and emotional feature factors, this article gives a definition of the TFIDF algorithm that integrates multiple features.

$$W_{ij} = \left(tf_{ij} \times idf_{ij} + W_{\text{span}_j} \right) \times W_{10c_{ij}} \times W_{\text{speech}_{ij}} \times M_{ij}. \tag{12}$$

The effect of user demand extraction is composed of three evaluation indicators. The experiment uses fit, recall, and $F1$ value ($F1$-measures) to evaluate the effectiveness of keyword extraction. The calculation method is shown in formulas (13)–(15).

$$P = \frac{C_{\text{correct}}}{C_{\text{extract}}}, \tag{13}$$

$$R = \frac{C_{\text{correct}}}{C_{\text{standard}}}, \tag{14}$$

$$F_1 - \text{Measure} = \frac{2pr}{p + r}. \tag{15}$$

To verify the effectiveness of the fusion multifunctional TFIDF method proposed in this paper and the method of retrieving unregistered vocabulary, the PUBLIC-PRAISE dataset is used to test the experimental results of various vocabulary function combinations, and the specific experimental results are obtained using the experimental evaluation indicators, as shown in the Table 1 shown.

The visualization result is shown in Figure 2.

To improve the performance of the keyword extraction effect of the method in this paper, the number of extraction of different keywords is set to $K = 5, 10, 15, 20, 25, 30$ for exploration, and the experimental results are shown in Figure 3.

The definition of bilinear upsampling is given in equations (16)–(18).

$$\begin{cases} f(R_1) = \dfrac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \dfrac{x - x_1}{x_2 - x_1} f(Q_{21}), \\[2ex] f(R_2) = \dfrac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \dfrac{x - x_1}{x_2 - x_1} f(Q_{22}), \end{cases} \tag{16}$$

$$f(P) = \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2), \tag{17}$$

$$\begin{aligned} f(x, y) = {} & \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y_2 - y) + \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y_2 - y) \\ & + \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y - y_1) + \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y - y_1). \end{aligned} \tag{18}$$

The optimized image I_c^* is initialized to $I_c^* = I_c$ to facilitate subsequent optimization work and avoid local minima. To calculate the loss value of I_c^*, I_s, and I_c, the definition of the loss function is shown as follows:

$$L_{\text{loss}} = \text{argmin} E_s\left(\phi(I_c^*), \phi(I_s)\right) + \alpha_1 E_c\left(\phi(I_c^*), \phi(I_c)\right) + \alpha_2 \gamma(I_c^*). \tag{19}$$

E_s is the square of the Euclidean distance ($L2$ normal distance) between the optimized image layer and the corresponding feature layer of I_c, i.e., the Markovian loss function defined as follows:

$$E_s\left(\phi(I_c^*), \phi(I_s)\right) = \sum_{i=1}^{m} \psi_i\left(\phi(I_c^*)\right) - \psi_{N(i)}\left(\phi(I_s)\right)^2. \tag{20}$$

The calculation method of the optimal feature block is as follows:

$$N(i) = \text{argmin} \frac{\psi_i\left(\phi(I_c^*)\right) \cdot \psi_j\left(\phi(I_s)\right)}{\left|\psi_i\left(\phi(I_c^*)\right)\right| \cdot \left|\psi_j\left(\phi(I_s)\right)\right|}. \tag{21}$$

To match the semantic label obtained by the semantic segmentation network with the result image of style transfer, the process uses the Hadamard image matrix to calculate the pixels between the semantic label and the result image. The calculation method is as follows:

$$R = Y \cdot Z. \tag{22}$$

The background pixel retention rate is as follows:

$$I = \frac{N}{M} \times 100\%. \tag{23}$$

The structure is similar.

TABLE 1: Comparison of TFIDF effects of different feature combinations.

Method	Precision	Recall	Fl-measure
TFIDF	0.353	0.400	0.376
TFIDF + unregistered vocabulary	0.378	0.426	0.400
TFIDF + unregistered vocabulary + vocabulary features	0.518	0.592	0.555
TFIDF + unregistered vocabulary + vocabulary features + emotional features	0.528	0.601	0.560
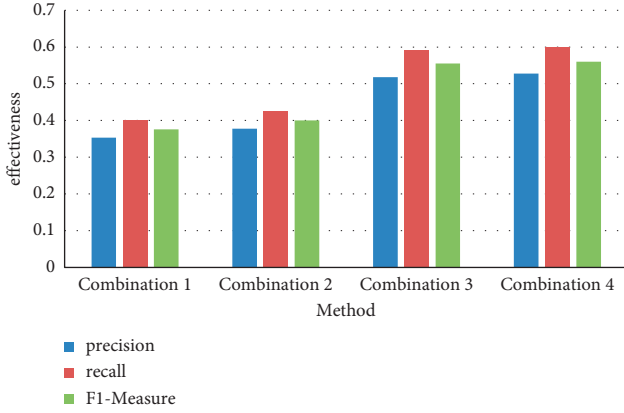
FIGURE 2: Comparison of TFIDF effects of different feature combinations.
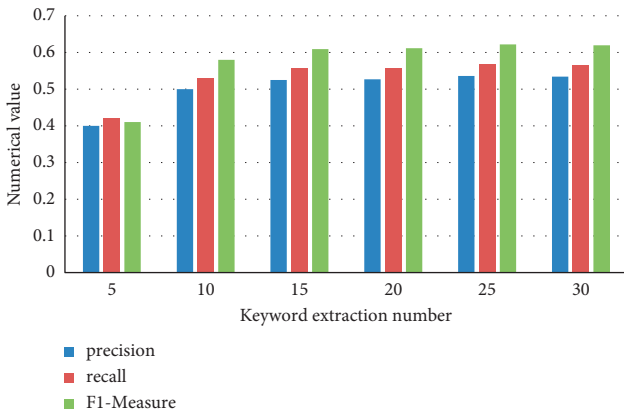
FIGURE 3: Comparison of the number of different keywords extracted.

$$\text{SSIM}(f,g) = L(f,g)gC(f,g)gS(f,g)$$

$$
\begin{cases}
L(f,g) = \dfrac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}, \\[2mm]
C(f,g) = \dfrac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2}, \\[2mm]
C(f,g) = \dfrac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_s^2 + C_2}.
\end{cases}
\tag{24}
$$

4. Financial Market Hedging Strategy and Optimization Research

4.1. Concept and Process. There are many types of hedging strategies, the most basic of which are long hedging and short hedging. The so-called long hedging is to buy derivative contracts at the moment when the spot price is expected to rise. The basis is that the derivatives and spot prices change in the same direction. When the spot price rises in the future, profits in the derivatives market are used to offset the spot market. Through this process, the purpose of hedging is achieved; the price forecast of short hedging is opposite to long hedging. At present, we mainly sell derivative contracts, however, the principle is the same as that of long hedging. They all use price fluctuations in the same direction. The principle is to achieve the purpose of using the derivatives market to avoid the price risk of the spot market.

In addition, there are initial assumptions in the initial selection and operation of the hedging strategy. If the hedger wants to hedge, he must conduct it in the spot risk exposure, and he can find a suitable time and amount of equivalent derivative products for his subject matter.

However, in the relationship between the actual spot market and the derivatives market, it is difficult to find derivatives that accurately correspond to the spot risk exposure, or there is a problem that there are no hedging instruments available in the derivatives market. In other words, the hedger needs to consider a cross-hedging strategy and should use another financial derivative that has a price correlation with the corresponding financial derivative.

The current real hedging in the derivatives market has also extended some more complex strategies based on long hedging and short hedging strategies. There are "rolling hedging" and "stripe hedging" in the short-term interest rate futures market. However, the core of these strategies is, in fact, still long and short strategies, which are only special forms extended under special circumstances. These expansion strategies have high requirements for use occasions. Hence, they can only solve the problems that exist in some special environments.

In the process of designing the actual hedging strategy, in addition to the principles of the derivatives market, it is also necessary to analyze the basic situation of the company itself. A more important aspect is the scope of hedging and the total amount of hedging. The scope of hedging refers to the company's outsourcing of raw materials, raw material production, raw material inventory, and spot products involved in product sales for hedging business operations through the corresponding futures varieties of the futures exchange.
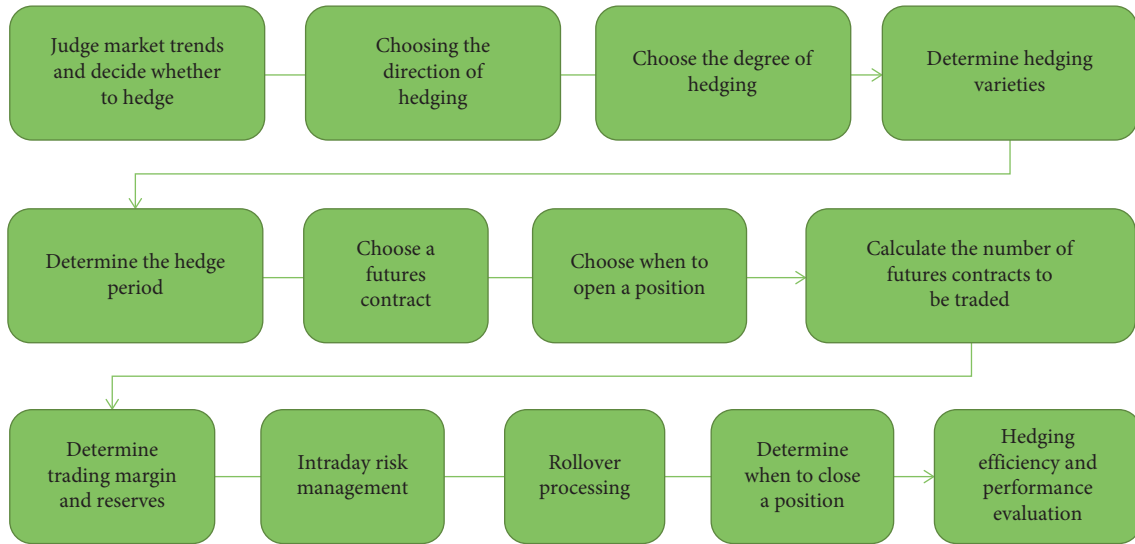
Figure 4: Hedging process.

The total amount of hedging includes the following parts: (1) the total amount of various specific types of hedging is limited to the company's total production of raw materials, total purchases of raw materials, and total product sales. (2) The total amount of various hedging operations shall not exceed the total amount limited by this plan; (3) When the hedging working group develops a hedging plan for specific project types, it will quantify step-by-step and implement step-by-step according to the actual situation and establish an optimized operating plan to allow the company's hedging. The period of preservation reaches the best state, as shown in Figure 4.

4.2. Strategy Classification and Evaluation. It is possible to conduct risk exposure analysis on the company's future raw material production, raw material procurement, product sales, and other business links to avoid the risks that market price fluctuations bring to the company's business.

The so-called unilaterally exposed enterprises are those with closed upstream and open downstream. On the contrary, the upstream and downstream enterprises are closed.

For companies with one-sided exposure, the focus of value preservation is the absolute price, and the price is unilaterally closed. Whether it is upstream or downstream, what the company has to do is to buy at the lowest possible price or sell at the highest possible price.

For this type of company, the basic method of hedging is relatively simple, either buying or selling, because one side of the price is fixed.

For companies with one-sided exposure, the uncertainty they face lies in the choice of projects, the choice of futures, the choice of the number of positions, and the choice of derivatives.

Therefore, the steps for a unilaterally exposed company to use the derivatives market for hedging can be mainly divided into three steps, which are as follows:

Firstly, determine the company's hedging target, then carry out price analysis and forecasting, and finally determine the hedging mode.

Different from unilateral exposure, bilateral exposure means that both the input price (cost) and output price change according to the price of the hedged object. This type of company usually comes from the middle reaches of the industrial chain, and the company's main business is trade or processing direction.

If the spot market price of the input and output products is the same, the company's core profit point is the difference between the processing fee and processing costs. In reality, the actual profit risk of the company lies in the price difference between the output product and the spot price of the input.

Since the actual spot target price usually fluctuates greatly, the value goal of the bilateral exposure company is to avoid the risk of the difference in the spot price of the sales item. In the actual hedging business, these companies are interested in the relative price of the spot rather than the absolute price of the spot target. The main purpose of the hedging plan is the hedging of the relative price.

According to the definition of hedging, an effective hedging operation must ensure that the combined benefits of the spot interest and the derivative market interest are positive, at least to ensure that the operation of the derivatives market allows the spot market to reduce losses.

In the current market, many companies do not know how to use hedging tools, especially after some state-owned enterprises have been involved in the "hedging" whirlpool of foreign futures markets. There are many doubts and negative comments in the market, which has led to the dismissal of companies that conduct hedging operations through the derivatives market.

In the actual operation of hedging companies, the evaluation of hedging effectiveness can be misleading. They may focus on the book profits and losses to evaluate the overall hedging effect. Obviously, this method is one-sided.

In addition, based on the correct hedging concept, the company can flexibly manage hedging based on actual risk exposure and cash flow in actual operations. It is not limited to copying and imitating the relatively successful models of other companies. Each company has its own special parts in terms of business conditions and financing methods, which need to be analyzed based on specific conditions, and different hedging models can be designed for different companies to better manage risk.

No matter how mature the company is, there is no guarantee that the actual hedging business will be perfect. There are uncertain factors, such as fundamental differences, cross hedging, and hedging costs. In addition to the company's natural profit-seeking nature, hedging operations are also the product of the game between controllable risk and excess profit, the product of the game between controllable risk and the maximum avoidance, and the reduction of basic differences and spread losses.

Therefore, when evaluating a company's hedging performance, firstly, determine the company's hedging target based on its own circumstances. In other words, to avoid spot market risks, the evaluation of the hedging effect should be based on whether to reduce risks and losses and how to operate the futures market. Loss can be reduced. Even if the risk exposure is not fully covered, hedging should be considered effective.

4.3. Effect Analysis. Table 2 shows the specific hedging strategies.

In early March 2020, according to the average price of nonferrous copper, the spot price of nonferrous copper was 45,390 yuan/ton. By May 2020, the spot price of nonferrous copper has fallen to 42,930 yuan/ton. Without hedging, a company's net asset loss would be RMB 12.3 million.

The specific calculation is as follows: a company's spot market profit and loss = (May-March copper market price) × company copper stock = (42930 yuan/ton − 45390 yuan/ton) × 5000 tons = 12.3 million yuan.

According to the aforementioned Shanghai copper futures hedging strategy, a company sold a total of 401 CU2005 and CU2006 contracts in early March. According to the set opening price, the average selling price in May was 44,601 yuan/ton, and the buying price was 42,861 yuan/ton and 42,721 yuan/ton.

According to calculations, the profit and loss of using the CU2005 contract is (44601 − 42861) × 2001 = 3.482 million yuan, the profit and loss when using the CU2006 contract is (44601 − 42721) × 2001 = 3.762 million yuan, the cumulative profit and loss is 7.244 million yuan, and the overall profit and loss situation is as shown in Table 3.

Taking the company's cash and cash equivalents as an indicator of the daily available cash flow, whether it is using Shanghai copper futures or Shanghai copper options for hedging, the funds occupied are covered by the company's daily operating cash flow, as shown in Figure 5. Cash flow pressure is relatively low.

According to the previous calculations, the hedging results, in this case, can be visually displayed, as shown in Table 4.

TABLE 2: Hedging content of enterprises using Shanghai copper futures.

Item	Content
Hedging purpose	Hedging the risk of 4,000 tons of standing inventory depreciation
Contract selection	Shanghai copper 2005 and 2006 contracts
Transaction direction	Sell
Length of preservation	March–May
Opening date	March Closing date May
Opening price range	44,000–45,000
Number of contracts	800 hands
Method of opening a position	Set up positions in batches, set up and sell 80 lots of Shanghai copper 2005 and 2006 contracts at 44,200, 44,400, 44,600, 44,800, and 45,000, respectively
Program features	Avoid systemic risks to a certain extent, and retain the possibility of copper prices going up

TABLE 3: The hedging effect of using Shanghai copper futures.

	CU2005	CU2006
Average selling price (yuan/ton)	44,601	44,601
Purchase price (yuan/ton)	42,861	42,721
Total number of contracts (hands)	401	401
Cover the risk part (tons)	2001	2001
Contract profit and loss (ten thousand yuan)	348.2	376.2
Total futures profit and loss (ten thousand yuan)	724.4	
Spot profit and loss (ten thousand yuan)	−1230	
Total profit and loss (ten thousand yuan)	−505.6	
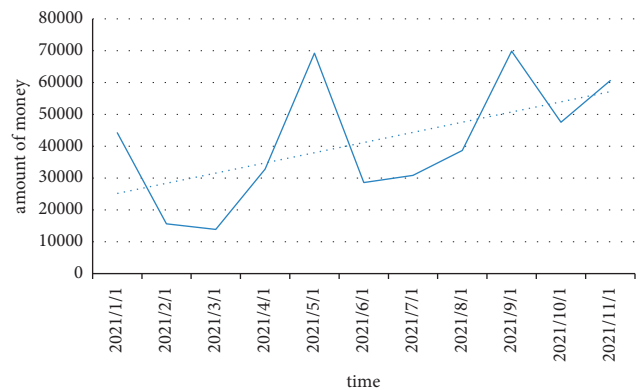
FIGURE 5: Corporate cash flow.

4.4. Optimized Plan Design. Shanghai copper options were listed for trading in September 2020. Like other commodity futures options, the basis of copper options contracts is copper futures contracts. Generally speaking, copper options are actually an option, i.e., traders can choose whether to

TABLE 4: Comparison of the results of no hedging and hedging.

—	No hedging	Shanghai copper futures
Total profit and loss (ten thousand yuan)	−1230	−505.6
Capital occupation (ten thousand yuan)	0	C. 901

withdraw their purchases at any time or the right to sell copper futures.

The general idea of using Shanghai copper options is similar to that of using Shanghai copper futures, both of which hedge the risk of falling copper prices while maintaining the potential for copper prices to rise. Therefore, the strategy of buying put options is adopted, the exercise price is set to 44,000–45,000, and the Shanghai copper options contract with the exercise price of 44,000 is finally selected, namely CU2005P44000.

Because of the low daily trading volume of Shanghai copper options, each order is limited to 100 lots, and the Shanghai copper options also use the bulk subscription method, with a total of 800 put options purchased.

When designing the hedging plan, the use of Shanghai copper options for hedging is another option for the company. It is mainly because Shanghai copper options use a lower option fee to obtain the future and buy and sell according to higher established prices. In the actual hedging process, because of the large trading volume and the difficulty of trading in the Shanghai copper options market, this option was not selected. However, in the following procedures, the benefits of this option can be analyzed through simulation, which can provide reference experience for the following design.

Because of the low daily trading volume of Shanghai copper options, the simulated backtesting process is difficult to meet the needs of one-time trading to cover the risk exposure. Hence, the transaction is carried out in batches, and the order is finally placed through 6 transactions. For the part of the transaction of 100 lots or more, the option premium is the average transaction price of the day.

According to the Shanghai copper option hedging strategy in the previous section, at the beginning of March, a company bought 800 lots of CU2005 P44000 at an exercise price of 44,000 yuan/ton. After the futures exercise, the partial profit and loss = (exercise price-CU2005 settlement on the exercise date) Price) × 4000 tons = (44000 yuan/ton − 41,725 yuan/ton) × 4000 tons = 9.1 million yuan, cumulative profit and loss = futures partial profit and loss-spot partial loss-option price = 9.1 million yuan − 12.31 million yuan − 428,800 yuan = −363.88 million yuan, the intuitive cost, and profit calculation are shown in Table 5.

The hedging optimization ideas in this paper have two main points: one is to use only Shanghai copper options and to introduce options that optimize the combination of Shanghai copper futures and options. The other is to consider whether the company can adopt the possibility of long-term hedging.

Although the optimization solution solved some of the problems in the case, it still exposed some risk points, which may require special attention in future operations.

First of all, in this case and optimization plan, the possibility of adjusting the production plan of the company was not considered. The company is at the downstream and end of the copper industry chain. The theoretical risk exposure should come from the upstream. Because of the adjustment of the production plan, the company's inventory is insufficient. To support daily production, the direction of hedging may be completely reversed. With the continuation of the current hedging direction, companies will also face losses.

In addition, because of the short hedging period of the company, neither the case nor the optimization plan has the expiry date of the derivative market contract, nor does it consider the risk of month-swapping and long-term coverage of the risk exposure. It is the hedging of other companies. Questions that cannot be skipped at the time. By comparing optimization plans and cases, we can understand the importance of derivatives trading volume in facilitating transactions, which also has a significant impact on the final hedging effect. If the trading volume in the options market is insufficient, the optimization plan has not been fully analyzed to support the possible consequences of hedging needs.

Neither the case nor the optimization plan considers other costs of the company's hedging, hedging expenses, other costs that need to be outsourced to specialized agencies, and other risks.

In addition, the company itself does not have a lot of hedging experience. Hence, it needs to seek the guidance of futures companies. There is a risk of leakage of the company's production and management secrets. Even if the company wants to carry out long-term hedging operations, it still needs training or employment and professional derivative personnel to design and manage its own hedging operations.

4.5. Analysis of Hedging Strategies in the Financial Market. In the process of using the derivatives market for risk management, companies establish a correct concept and make it clear that the original intention of the company to use derivatives is to avoid risks and not to profit through speculation. It requires a clear understanding of the risks that the company can bear. In the futures market, the high leverage provided by margin trading allows companies to support larger transactions with less capital. At the same time, it also introduces new risks and has higher requirements on whether companies understand the nature of futures market transactions.

Enterprise experience shows that the direction of the company's hedging may not match the company's position in the industry chain. The company is located in the downstream or terminal position of the copper industry

TABLE 5: Hedging backtest results using Shanghai copper options.

CU2005P44000 transaction situation	
Option fee (yuan/hand)	Volume (hands)
501	201
336	101
550	201
601	201
651	101
Analysis of profit and loss results	
Total option premium (ten thousand yuan)	42.88
Partial profit and loss of futures after exercise (ten thousand yuan)	910
Spot profit and loss (ten thousand yuan)	−1231
Accumulated profit and loss (ten thousand yuan)	−363.88

chain. Traditional hedging thinking requires companies to hedge upstream risk exposures and is more sensitive to copper price increases. However, the company's inventory situation, in fact, is very good. For the company's daily production, it is more urgent to prevent inventory depreciation. In short, the company's actual risk exposure is similar to that of copper material manufacturers. What is worrying is that the price of copper has fallen. Therefore, the direction of hedging is opposite to the experience of hedging in the industry chain.

The derivatives market is the main place for companies to hedge, and it has an irreplaceable effect on companies to avoid market risks. However, it is undeniable that in the process of using the derivatives market, future market price fluctuations, margin risks, and liquidity risk may arise. Failure to correctly understand the existence of these risks will reduce the effectiveness of hedging and increase losses for companies.

As mentioned above, companies need to improve their professional knowledge and correctly understand the risks in the derivatives market. In addition to looking for specialized agencies to guide hedging, companies themselves also need to improve the professional quality of internal risk management institutions, fully understand the futures market, and improve the futures market. The risk identification and management functions enable the company's employees engaged in related businesses to strengthen their understanding of the futures market and strengthen the risk concept of all employees.

In addition, it should be recognized that under the current market environment and accumulated experience, risks in the derivatives market can be controlled by effective means, rather than allowing companies to withdraw.

Whether using futures or hedging options, capital occupation is inseparable. In the futures market, capital occupation is mainly reflected in the margin, while in the options market, capital occupation is mainly reflected in the option premium. Generally speaking, option costs are relatively fixed. Hence, the option use of a hedge fund can be determined at the time of trading. The futures market is traded on margin, and the amount of margin is much lower than the actual transaction amount supported. However, for companies with large positions, margin occupancy should also be considered.

5. Conclusion

The automobile industry is one of the important pillar industries of China's economy and has made great contributions to the vigorous development of China's economy. However, the future development of traditional automobiles has many limitations, and the new energy and green automobile industry are thriving with huge development potential, providing a strong driving force for promoting economic development. As new energy green vehicles are more energy-saving and environmentally friendly, the state and local governments provide fiscal and tax subsidies for the development of new energy vehicles, which brings many opportunities for the development of new energy vehicle intelligent manufacturing. This article starts with the company's actual hedging needs, analyzes the company's risk exposure, determines the hedging target, analyzes the copper spot market fundamentals, and predicts the price trend. In the process of simulating the company's hedging, we mainly evaluate the effect of hedging in terms of operability, current market profit and loss, and capital occupation. After evaluating and analyzing the case, this paper proposes an optimization plan based on the existing problems and conducts an evaluation of the backtest results.

Data Availability

The data used to support the findings of this study are available from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] H. Bloch, S. Rafiq, and R. Salim, "Economic growth with coal, oil and renewable energy consumption in China: prospects for fuel substitution," *Economic Modelling*, vol. 44, pp. 104–115, 2015.

[2] E. D. Attanasi, "Bitumen prices and structural changes in north American crude oil markets," *Natural Resources Research*, vol. 25, no. 4, pp. 1–10, 2016.

[3] S. J. Byun, "Speculation in commodity futures markets, inventories and the price of crude oil," *Energy Journal*, vol. 38, no. 5, pp. 93–113, 2017.

[4] Q. Zhang, *Dispersion of Vehicle Exhaust Gas with Numerical Simulation CFD*Nanjing University of Science and Technology, Nanjing, China, 2007.

[5] Z. Ning, C. Cheung, Y. Lu, M. Liu, and W. Hung, "Experimental and numerical study of the dispersion of motor vehicle pollutants under idle condition," *Atmospheric Environment*, vol. 39, no. 40, pp. 7880–7893, 2005.

[6] X. L. Wu, "Environmental policy, anti-pollution effort degree and ecological environment quality-simulation analysis based on three-sector DSGE model," *Collected Essays on Finance and Economics*, vol. 4, pp. 101–112, 2017.

[7] Ipcc, *Special Report: Global Warming of 1.5°C*, IPCC, Geneva, Switzerland, 2018.

[8] A. G. Olabi, T. Wilberforce, and M. A. Abdelkareem, "Fuel cell application in the automotive industry and future perspective," *Energy*, vol. 214, Article ID 118955, 2021.

[9] C. Feenstra Robert and S.-j. Wei, *Introduction to China's Growing Role in World Trade*National Bureau of Economic Research, Massachusetts, MA, USA, 2009.

[10] Th. D. Willett, P. Liang, and N. Zhang, "Global contagion and the decoupling debate," in *The Evolving Role of Asia in Global Finance (Frontiers of Economics and Globalization, Volume 9)*, Y.-W. Cheung, V. Kakkar, and G. Ma, Eds., Emerald Group Publishing, Bingley, West Yorkshire, 2011.

[11] D. Schmidt, "The financial crisis and its impact on China," *China Analysis Research Group on the Political Economy of China*, Trier University, Trier, Germany, 2009.

[12] Z. Chen, "A massive deflationary shock to the world," in *A Symposium of Views: If the Chinese Bubble Bursts*International Economy, China, 2010.

[13] A. Strutt and T. L. Walmsley, "Trade and sectoral impacts of the financial crisis: a dynamic cge analysis," in *Proceedings of the ARTNet Asia-Pacific Trade Economists': Trade-Led Growth in Times of Crisis*, Bangkok, Thailand, November 2009.

[14] D. Yang, E. Qi, and Y. Li, "Quick response and supply chain structure with strategic consumers," *Omega*, vol. 52, pp. 1–14, 2015.

[15] D. G. Hobson, "Robust hedging of the lookback option," *Finance and Stochastics*, vol. 2, no. 4, pp. 329–347, 1998.

[16] S. Figlewski, "Options arbitrage in imperfect markets," *The Journal of Finance*, vol. 44, no. 5, pp. 1289–1311, 1989.

[17] N. H. Bingham and R. Kiesel, *Risk-Neutral Valuation, Pricing and Hedging of Financial Derivatives*, Springer, London, UK, 1998.

[18] C. Albanese and S. Tompaidis, "Small transaction cost asymptotics and dynamic hedging," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1404–1414, 2008.

Research Article

Research on Innovation Ecosystem of Dairy Industry Cluster Based on Machine Learning and Improved Neural Network

Yang Hui [iD]**, Yanni Jiao, Can Cui, and Ke Ma**

College of Economics and Management, Northeast Agricultural University, Heilongjiang, Harbin, China

Correspondence should be addressed to Yang Hui; yanghui@neau.edu.cn

With the rapid development of global economy, the industrial cluster has become the new trend of world economic development. So far, the industrial cluster mode with space as the main division has been formed. The cooperation of industrial clusters is dynamic. The industries in the cluster cooperate with each other for mutual benefit and win-win, occupying a place in the fierce market competition. In addition, the industrial cluster is also conducive to strengthen the international economic division of labor and points out the direction of regional industrial transfer. China is now in a critical period of economic development; the industrial cluster plays an important role in China's industrial transformation and economic development. At present, the most common regional development mode in China is industrial cluster. The emergence of industrial cluster accelerates the development of industrial regional economy and the balance of industrial layout. Industrial clusters occupy a key position in China's economic industry chain, and the development and change in industrial clusters will directly affect the development of the entire industrial chain. This study first simulates the evolution path of industrial cluster and then establishes the relevant data model. Finally, through repast simulation, it puts forward conclusions and suggestions according to the verification results. The construction of dynamic model can realize the simulation of industrial cluster theory. According to the simulation results, we can find that the ecosystem in different stages will produce different characteristics; the formation and evolution of industrial clusters are actually the epitome of market development. In this process, the government guidance and market regulation are needed to accelerate the formation of industrial cluster ecosystem and increase the scale of industrial cluster.

1. Introduction

After the reform and opening up, China's industrial clusters have developed rapidly. On the one hand, the economic development after the reform and opening up relies more on the development of regional economy, and the development of industries in the same region is basically the same, which is conducive to the formation of industrial clusters; on the other hand, there is a highly specialized division of labor and close industrial ties between different industries. It has laid a rich foundation for the formation of industrial clusters [1]. At present, China has formed a relatively mature industrial cluster ecosystem. Different industries cooperate with each other and occupy a dominant position in the market. In the development process of China's industrial clusters, each stage has different changes. Industrial clusters should also make corresponding changes according to the characteristics of different stages, follow the trend of market development, and maintain the advantages of industrial clusters. So far, the common idea in academic circles is that China's industrial clusters will develop in the direction of cluster ecosystem, thus forming the advantage of market competition [2]. However, the current academic research on the cluster ecosystem is not deep enough, and the social understanding of the cluster ecosystem only stays on the concept of "ecolgical rent," but there is no accurate understanding of the specific concepts of cluster ecosystem and "ecological rent." This is the focus of our future research; we need to popularize the concept of cluster ecosystem to the whole society.

The theory of multi-industry cluster and social organization is a kind of complex system. Among the organizations and institutions that make up the cluster, microenterprises occupy the most important position. The formation of industrial clusters does not depend on the drive of a single enterprise, but requires the collaborative development of enterprises in the industry [3]. To form a mature and orderly industrial cluster structure, we should constantly innovate and update the cluster system and related decisions to ensure that the industrial cluster can keep up with the pace of the times. It is also very important for industrial clusters to make correct decisions. Good decisions can promote the rapid development of industrial clusters, even change the macrostate of industrial clusters, and promote the sustainable development of industries. Therefore, before carrying out the policy, we should fully consider whether the policy meets the actual requirements and whether the policy is rigorous enough. Only after the rationality of the policy is fully determined, the policy can be issued. In addition, through the establishment of industrial cluster simulation model to test whether the various factors affecting the development of industrial cluster are reasonable.

2. Related Work

The idea of industrial cluster was first put forward by Adam Smith. Literature [4] thinks that division of labor is very important in industrial cluster. Only by determining the role of each part of industrial cluster, we can give full play to the advantages of industrial cluster and form scale economy. The development of industrial clusters should not only pay attention to the cooperation between various departments but also pay attention to the refinement of functions. On the research of industrial clusters, the theory of industrial location proposed by literature [5] and the theory of industrial location proposed by literature [6] have become the important ideas for the development of industrial clusters, and more and more studies have been carried out with these two ideas. However, foreign scholars only focus on the causes of industrial clusters and do not conduct in-depth research on the follow-up development of industrial clusters. Research: in addition, most of the foreign research on industrial cluster is based on the theory of industrial cluster. Literature [7] and others began to study the theory of technology spillover, and literature [8] carried out the research on knowledge management of industrial clusters. In the research, we should first study different cluster organizations separately. Literature [9] classifies the research focus as the knowledge resources of industrial clusters. At present, the research focus of industrial clusters is the research integration between different schools of industrial clusters and the theoretical knowledge in industrial clusters. This change shows that the academic research on industrial clusters has begun to transition from theory to practice, which is a qualitative leap for industrial clusters.

Literature [10] puts forward that in the whole industrial cluster ecosystem, some ecological factors related to the industrial cluster ecosystem jointly promote the development and progress of the cluster industry, promote the formation of the cluster effect within the industry, and promote the cooperation of all parts of the cluster. In reference [11], the industrial cluster ecosystem should not be a simple industrial collection, it is formed in the long-term evolution process, and the regional economic ecosystem is a dynamic system. In the development of industrial clusters, we should pay attention to the establishment of industrial brands to form competitive advantages. Literature [12] believes that the emergence of brand effect is with the development of industrial clusters, and when the industrial clusters develop to a certain extent, it will promote the formation of brands within the industry, thus producing brand effects externally. Literature [13] has proved that for industrial clusters, transformation is the only way, and the key to achieve transformation is to maintain the competitive advantage within the cluster. Literature [14] analyzes the impact of network on industrial clusters from the perspective of network. The analysis results show that the use of network in the cluster can strengthen the relationship among various parts of the cluster and can also improve the competitive advantage of individual enterprises in the cluster. According to the resource-based theory proposed in reference [15], heterogeneous resources of enterprises can be divided into tangible and intangible resources. Although the existence forms of these two resources are different, they are indispensable resources in the process of cluster. These resources are the source of competitiveness and creativity of enterprises and play an irreplaceable role. Literature [16] believes that rich resources can ensure the dominant position of enterprises in the competition. In literature [17] for the enterprises in the cluster, it is very important whether they have their own rooted resources, which is the key element to distinguish enterprises from other enterprises in the cluster, and the core and key in the cluster ecosystem. Literature [18] has proved the importance of heterogeneous resources to enterprises. Heterogeneous resources can catalyze the emergence of competitive advantages of enterprises in clusters. In addition to promoting the development of individual enterprises, they can also promote the common development of industrial clusters and even the entire industrial chain. According to the literature [19], competition and cooperation are the two basic ways of industrial cluster formation. The existence of these two ways ensures the continuous creation of new things within the industrial cluster, with continuous vitality and sustainable development. The industrial cluster ecosystem is similar to the natural ecosystem, but the competition between enterprises cannot be compared with the relationship in the natural ecosystem. The competition between enterprises includes both struggle and cooperation, which is different from the single competition in nature. In fact, competition is the best cooperation. Based on the theory of cooperative competition, literature [20] confirmed that orderly competition in the cluster can promote the common development of enterprises and clusters and make the enterprises in the cluster improve their competitiveness in the market. Literature [21] holds that competition is necessary for cooperation. In fact, cooperation is developed from competition, and cooperation is a consensus reached in the process of competition. According to the niche theory, literature [22] holds that market competition is actually a struggle for innovation

ability, especially for enterprises of the same nature, and if they want to stand out among the same kind, they must have strong innovation ability and seek new profit growth points; in addition, competition is necessary for the development of clusters, and competition can promote cooperation and division of labor among enterprises. At the same time, it will promote the exchange and integration of different enterprises in the same industry chain. In reference [23], the industrial cluster ecosystem is a dynamic system. In the long-term development process, each component of the ecosystem realizes common survival and coevolution, aiming at improving the overall competitiveness and making common progress.

Agent theory was first proposed in the field of artificial intelligence. As mentioned in literature [24], so far, agent theory has not only existed in the field of artificial intelligence, but also applied to various fields of society. For different fields, the meaning of agent theory is different. Different fields will give different meanings to agent theory to meet the needs of this field. Therefore, it is very difficult to determine the unified definition of agent theory. Although different fields have different definitions of agent theory, each domain has some common characteristics, such as autonomy, intelligence, autonomy, and sociality.

3. Technology and Theory of the Dairy Industry Cluster Ecosystem

3.1. Modeling Method of the Multiagent System. Multiagent system comes from the field of artificial intelligence and has been applied to many fields. MAS is a system composed of multiple agents. MAS can achieve the task goal through the cooperation between agent systems. MAS is actually a complex joint system. This system can realize data simulation and solve the actual problems.

Before using multiagent system for research, we should first determine the specific role of each element in the system and clarify the relationship between different elements; secondly, we should extract the key elements of the system, establish the agent model of the system, and set the rule conditions of the agent model; finally, after the establishment of the agent model, the model is run and data statistics and analysis are carried out, and the analysis results with the actual results are compared.

In MAS system, the agent is defined as individual, and the relationship in MSA system can be regarded as the static relationship model between different individuals; in the construction of multiagent class, the interaction and cooperation are carried out through the establishment of each class to form a model for operation and solve each level and task in the system.

3.2. Quantitative Model of Competition Cooperation Interaction Game of Industrial Cluster

3.2.1. Basic Assumptions of the Model

(i) Suppose that there are only two enterprises (1, 2), and the two enterprises are substitutes. The total supply and demand function in the market is as follows:

$$Q = q_1 + q_2,$$
$$p(Q) = p(q_1 + q_2). \tag{1}$$

(i) P is the price, and Q(p) is the original demand function. The two companies here have the same unit cost c and total output Q. Suppose

$$p(Q) = A - Q,$$
$$C_i(Q) = q_i c. \tag{2}$$

(i) Profit R is paid by the enterprise itself, and it is a function of output Q.

3.2.2. Game Model Analysis. We can get the model we want by comparing the companies in a single time. According to the model obtained, the profit of the two companies is calculated as follows:

$$R_1 = R_2 = Q * P = \frac{A-c}{4} \times (P-c) \frac{(A-c)^2}{8}. \tag{3}$$

The profit of a single enterprise is increased, and the scale is expanded. The profit obtained is as follows:

$$R_1 = q_1(p-c) = q_1(A - Q - c) = q_1[A - (q_1 + q_2) - c]. \tag{4}$$

The optimized first-order condition is as follows:

$$\frac{\partial R_1}{\partial R_2} = A - (2q_1 + q_2) - c = 0. \tag{5}$$

The relationship between the profit and output of the company can be obtained. In the case of other companies also cooperating,

$$q_2 = \frac{A-c}{4},$$
$$q_1 = \frac{3(A-c)}{8}. \tag{6}$$

Thus,

$$R_1 = q_1(p-c) = q_1[A - (q_1 + q_2) - c] = \frac{9(A-c)^2}{64},$$
$$R_2 = q_2(p-c) = q_2[A - (q_1 + q_2) - c] = \frac{3(A-c)^2}{32}. \tag{7}$$

When another company chooses not to cooperate due to its own choice, it still follows the principle of profit maximization:

$$R_2 = q_2(p-c) = q_2[A - (q_1 + q_2) - c]. \tag{8}$$

The first-order optimal conditions are as follows:

$$\frac{\partial R_1}{\partial R_2} = A - (q_1 + 2q_2) - c = 0. \tag{9}$$

Comprehensively, the above two formulas are solved:

$$q_1 = q_2 = \frac{A - c}{3},$$

$$R_1 = R_2 = \frac{(A - c)^2}{9}. \tag{10}$$

Regarding the use of game analysis, if both parties choose to cooperate at this time, they will get the effect of cooperation.

$$K = \frac{(A - c)^2}{8}. \tag{11}$$

At the same time, for the intention of cooperation, if one party is willing to cooperate and the other party is unwilling to cooperate, for the party who does not want to cooperate:

$$L = \frac{9(A - c)^2}{64}. \tag{12}$$

Partners get

$$M = \frac{3(A - c)^2}{32}. \tag{13}$$

If both parties are unwilling to cooperate, then the result is as follows:

$$N = \frac{(A - c)^2}{9}. \tag{14}$$

It can be concluded that the order of benefits after cooperation between them is as follows: $L > K > N > M$.

If companies conduct multiple comparisons and evaluate the results from multiple aspects, this formula can be derived as follows:

$$K = R + \delta^2 R + \delta^3 R + \cdots + \delta^n R = \frac{R(1 - \delta^n)}{1 - \delta} = \frac{(A - c)^2}{8(1 - \delta)}. \tag{15}$$

The profits gained by the refusal of cooperation between enterprises are as follows:

$$N = \frac{(A - c)^2}{9(1 - \delta)}. \tag{16}$$

Therefore, only cooperation is the guarantee of long-term development. For this reason, competition and cooperation within industrial clusters are inevitable. The relationship between the various industries in the process of cooperation is shown in Figure 1.

As shown in Figure 1, the division of labor between different industries in the company is very clear. They let their employees do some major production and development work. The core technology and scientific research work are responsible for the development department. It will be completed by some other companies.

3.3. Mathematical Model of Industrial Cluster Learning

3.3.1. Model Establishment. There are two main aspects that have an impact on the industrial production function, namely the total amount of participation and the effective yield. According to the relationship between them, we can build the following functions:

$$Y = E(,) * F_p. \tag{17}$$

Different mathematical symbols in the function have different meanings. E and Y_p represent the effective yield and participation amount, respectively. k1 and k2, respectively, represent the private knowledge and learning knowledge in the production process, and k represents all the knowledge output, where Y and B represent the functional changes in the learning knowledge and private knowledge in the process, and their scope is γ, $\theta \in (0.1)$. Based on this relationship, we can build the following function:

$$E(,) = E(K_c, k_1; \gamma, \theta) = aK_c^{\gamma} k_1^{\theta}. \tag{18}$$

The sample size of the research object is set as x, and the total amount of learning knowledge, learning rate, and total enterprise knowledge from these samples are extracted as follows:

$$K_c = \int_0^k k_2 dt = xk_2,$$

$$\pi = k_2/k, \tag{19}$$

$$k = k_1 + k_2.$$

FR and σ are used to represent the mathematical characteristics of enterprise knowledge under different elements:

$$K_c = xk_2 = x\pi k = K_c^{\sigma} F^R,$$

$$K_c = \left(x\pi F^R\right)^1 / (1 - \sigma), \tag{20}$$

$$K_c^{\gamma} = \left(x\pi F_R\right)^{\gamma/1 - \sigma}.$$

Because

$$k_1 = k - k_2,$$

$$k_1^{\theta} = (k - k_2)^{\theta} = (1 - \pi k)^{\theta} k^{\theta},$$

$$k^{\theta} = \left(K_c^{\sigma} F_R\right)^{\sigma} = \left(x\pi F_R\right)^{\partial \sigma/1 - \sigma}, \tag{21}$$

$$k_1^{\theta} = (1 - \pi)^{\theta} k^{\theta} = (1 - \pi)^{\theta} \left(x\pi F_R\right)^{\partial \sigma/1 - \sigma}.$$

Combining equations (19)–(21), these three formulas can be solved:

$$E(,) = aK_c^{\gamma} k_1^{\theta} = a(1 - \pi)^{\theta} \left(x\pi F_R\right)^{\lambda},$$

$$\lambda = \frac{\lambda + \theta\sigma}{1 - \sigma}. \tag{22}$$

3.3.2. Model Analysis. Because the sample size is fixed, we can establish the following relationships according to the different uses of the input elements:
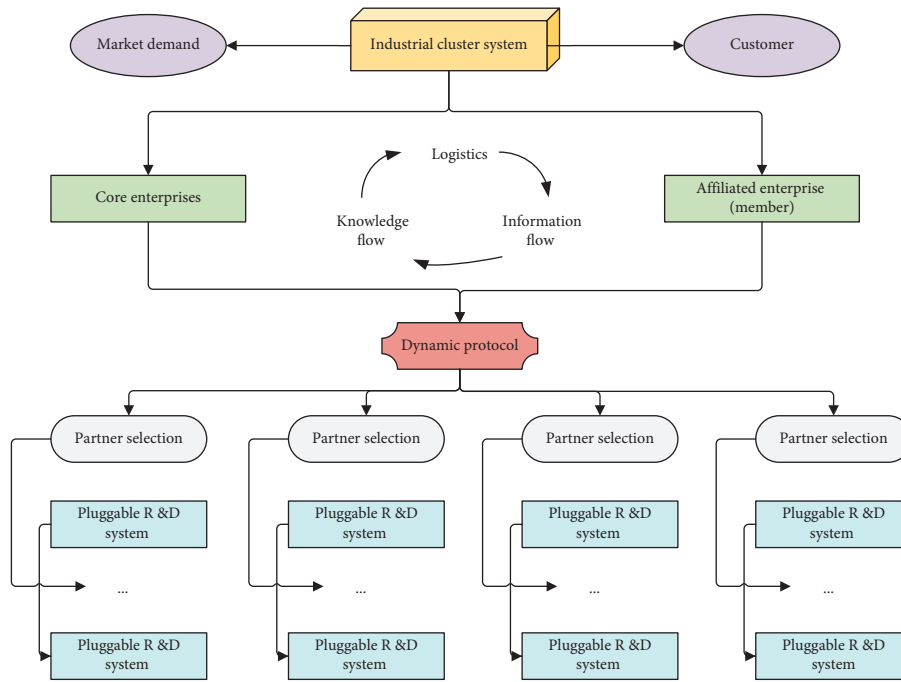
FIGURE 1: Mechanism of interfirm cooperation in the industrial cluster system.
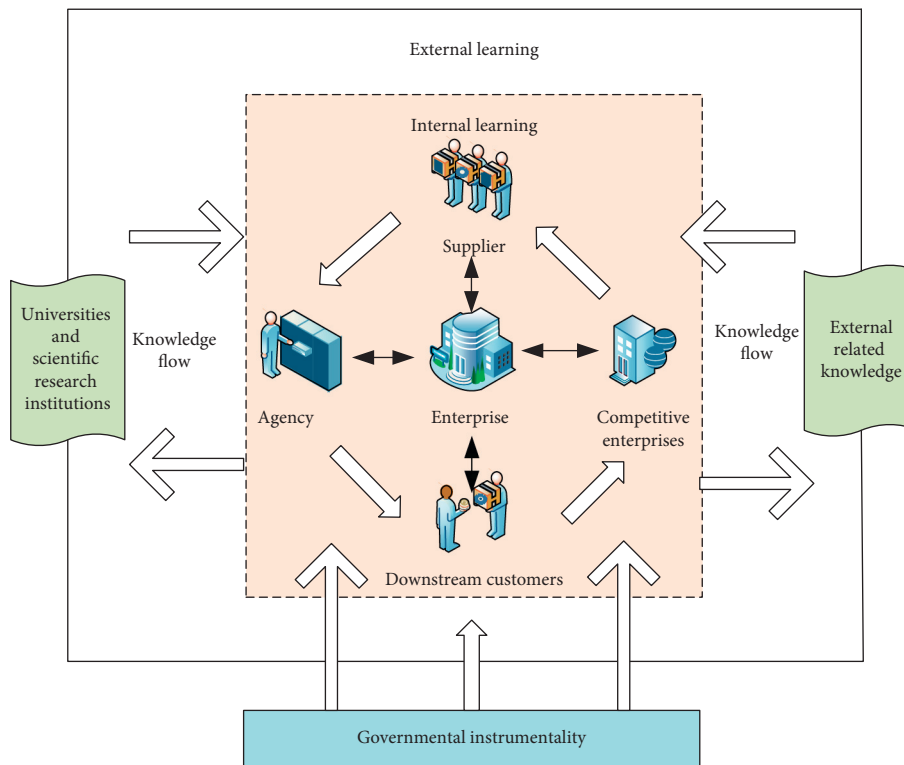
FIGURE 2: Schematic diagram of the industrial cluster learning system.

$$F = F_P + F_R,$$
$$y = a(1-\pi)^\theta (x\pi F_R)^\lambda F_R. \tag{23}$$

After establishing the Lagrangian function and solving,

$$L = y - \lambda[(F_P + F_R) - F],$$
$$\lambda F_P = F_R, F_R = F/(1+\lambda),$$
$$F_P = \lambda F/(1+\lambda). \tag{24}$$
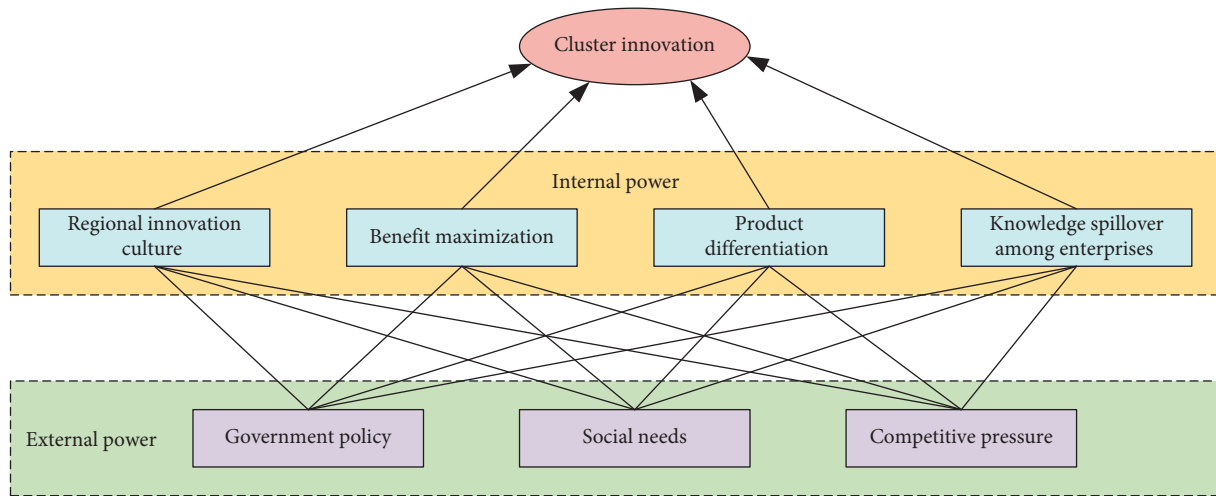
FIGURE 3: Composition diagram of the main body of industrial cluster innovation.

TABLE 1: Retail price index of Xinjiang dairy products from 1998 to 2009.

Category name	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Cumulative increase	Average increase
Dairy products	103.3	101.71	99.6	99.4	99.4	99.9	100.6	98.7	100.8	108.2	122.2	102.0	105.7	102.8
Pasteurized milk	104.9	100.8	100.3	99.7	99 5	100.4	100.3	99.s	100.1	108.4	125.6	101.8	105.7	103.2
Yogurt									103.7	112.4	111.0	101.6	105.7	107.1
Milk powder	98.1	104.4	96.5	97.7	96.2	97.3	99.6	101.4	105.3	102.6	116.3	105.5	105.7	101.6
Other				98.6	100.7	98.3	102.3	93.2	101.3	103.8	106.5	101.4	105.7	100.6

TABLE 2: Retail prices of liquid milk in various regions from 2011 to 2019.

Specification grade	2011	2012	2013	2014	2015	2016	2017	2018	2019
250 g bag	2.80	2.80	2.80	2.80	2.80	2.80	3.16	4.00	4.64
Bulk milk	2.50	2.50	2.50	2.50	2.50	2.50	2.62	3.20	3.20
Maiquer milk					7.50	7.50	7.50	9.40	9.40
Bulk milk	2.42	2.40	2.39	2.28	2.40	2.88	2.92	3.76	4.00
Bulk milk	2.00	2.00	2.06	2.10	2.10	2.11	2.60	3.00	2.97
Gary bag						5.50	5.50	6.00	7.30
Real estate bag	2.00	2.00	2.00	2.04	2.15	3.00	3.40	4.20	4.20
Ruiyuan bag	2.44	2.80	6.08	5.92	5.68	5.20	5.32	6.56	6.72
Bulk milk	2.00	2.00	2.00	2.00	2.00	2.00	2.66	3.50	3.22
Bulk milk	2.10	2.07	2.00	2.00	2.00	5.00	5.00	5.25	6.50
Bulk milk	2.40	2.40	2.25	2.00	2.80	3.12	3.12	5.00	5.48
Bulk milk	1.74	1.51	1.42	1.50	1.53	1.90	1.97	2.43	3.08
500 g bag	3.00	3.00	3.00	3.00	3.00	6.00	6.00	6.00	6.00
Maiquer milk	5.00	5.00	5.00	5.00	7.50	7.50	7.60	9.00	10.15
Bulk milk	1.51	1.55	1.63	2.00	1.80	1.90	2.21	3.04	2.88
Bulk milk	2.17	2.21	2.28	2.18	2.29	2.17	2.49	3.04	2.72
Bulk milk	2.00	2.00	2.00	2.00	2.00	2.08	2.58	3.00	3.03
Bulk milk	1.48	1.55	1.60	1.68	1.68	1.82	2.17	3.50	2.71
Bulk milk	0.60	0.60	1.50	1.50	2.00	2.04	2.61	3.32	2.54
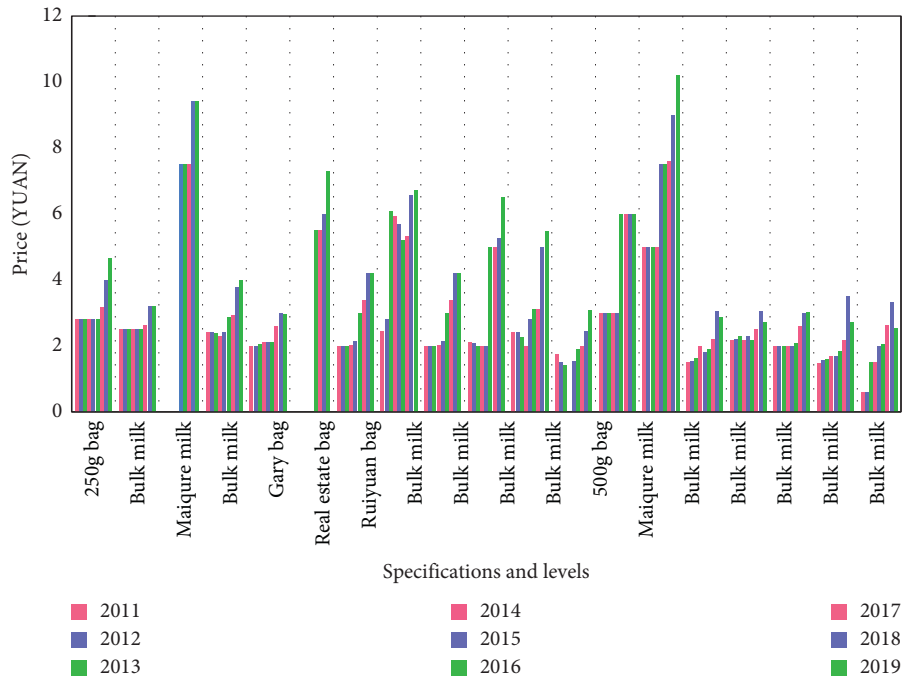
FIGURE 4: Changes in retail prices of liquid milk in various regions from 2011 to 2019.

TABLE 3: Per capita consumption expenditure of dairy products by urban residents in various provinces, municipalities, and districts across the country from 2015 to 2019.

Area	2015	2016	2017	2018	2019
National average	124.70	132.37	138.62	15023	160.72
Beijing	255.99	255.79	270.43	279.86	279.45
Tianjin	139.08	135.49	149.17	153.15	182.18
Hebei	117.51	128.16	134.25	140.61	150.19
Shanxi	128.45	14335	154.26	171.83	183.97
Inner Mongolia	109.24	111.99	120.14	124.96	13332
Liaoning	129.81	138_18	140.14	149.13	153.06
Jilin	86.59	102.47	105.30	106.45	120.13
Heilongjiang	101.37	10630	95.60	10633	112.49
Shanghai	241.82	250.03	246.88	267.34	313_04
Jiangsu	141.95	14422	156.29	173.07	18226
Zhejiang	139.17	143.51	155.28	16025	169.31
Female emblem	117.20	137.02	154.65	175.09	192.84
Fujian	153.00	16133	152.34	159.06	164.79
Jiangxi	94.30	104.51	119.63	129.09	157.46
Shandong	149.65	159.16	159.61	176.72	197.63
Henan	91.96	10623	115.26	12322	145.11
Hubei	90.27	97.69	110.25	124.17	143.62
Hunan	87.65	98.52	119.20	117.88	106.98
Guangdong	124.01	127.39	134.49	157.28	147.10
Guangxi	85.31	100.34	99.58	95.03	118.27
Hainan	68.98	85.10	69.53	83.86	104.07
Chongqing	178.21	183.91	168.05	187.78	181.94
Sichuan	130.10	131.07	138.86	146.25	161.74
Guizhou	82.48	86.59	96.40	107.11	118.30
Yunnan	67.73	75.79	78.23	68.82	56.58
Tibet	333.64	419.02	339.90	269.24	284.90
Shaanxi	103.81	112.54	12431	133.19	151.78
Gansu	115.57	125.97	136.80	137.93	140.87
Qinghai	96.29	100.00	104.44	112.77	138.73
Ningxia	103.70	116.89	124.17	135.05	154.82
Xinjiang	110.10	117.28	113.75	123.65	132.77

TABLE 4: Per capita consumption of liquid milk in major countries in the world.

Country	2014	2015	2016	2017	2018	2019
Denmark	133.10	134.00	135.70	136.70	137.20	135.70
France	92.50	92.60	92.50	92.90	91.70	92.20
Germany	89.90	90.10	91.00	93.90	91.90	92.70
Ireland	14420	157.00	157.00	155.70	148.60	
EU 25 countries		90.70	91.70	92.80	92.50	92.80
Netherlands	125.80	120.20	120.50	127.70	127.10	127.10
United Kingdom	114.80	111.80	111.40	111.60	108.90	111.20
Norway	123.40	121.90	121.30	115.20	115.60	114.80
Switzerland	92.50	101.20	106.30	108.50	110.40	111.90
Canada	95.80	94.90	91.80	95.00	95.60	94.50
Mexico	38.10	36.30	39.10	3 9.00	39.80	38.30
United States	87.60	86.30	85.50	84.90	84.50	83.70
Argentina	69.80	70.90	62.20	62.10	65.80	
Brazil	74.60	74.50	71.20	71.00	72.00	73.20
Australia	102.30	103.10	104.00	105.50	106.20	106.30
New Zealand	99.00	98.70	97.00	90.00	90.00	90.00
Japan	3 7.00	35.60	39.20	38.50		
China	1.00	2.20	3.20	5.10	7.20	8.80

TABLE 5: 2017–2019 per capita annual dairy consumption expenditure of urban households in a certain province with different income levels.

Income grouping	2017	2018	2019	Cumulative growth %
Total average (yuan)	113.75	123.65	132.77	8.0
Lowest income households	39.70	54.40	59.21	22.9
Among them, difficult households	23.40	38.75	46.07	42.2
Low-income households	71.58	79.46	101.41	19.3
Middle-lower	96.53	111.50	132.20	17.0
Middle-income households	115.13	120.81	136.41	8.9
Upper-middle class	139.23	144.66	155.05	5.5
High-income households	158.82	182.38	162.30	1.9
Highest income households	207.32	197.61	211.58	12

After substituting formula (23),

$$y = A(1-\pi)^{\theta}\lambda x^{\lambda}\pi^{\lambda}F_R^{1+\lambda},$$
$$A = a\lambda^{\lambda}(1+\lambda)^{-(1+\lambda)},$$
$$\text{if } F = F_P + F_R = 1, \quad (25)$$
$$y = f(\pi, x) = A\phi(\pi)\varphi(x),$$
$$\phi(\pi) = (1-\pi)^{\theta}\pi^{\lambda}, \varphi(x) = x^{\lambda}.$$

By analyzing the formula, we can draw the conclusion that there is a positive mathematical relationship between the output of the cooperative group and the knowledge learning rate and the number of enterprises, and the geographical location of the enterprise is also more positive than the output of the cooperative group. The output of the cooperative group mainly depends on the company's technology and knowledge characteristics.

3.4. Analysis of Innovation Motivation Mechanism of Industrial Clusters

3.4.1. Path Analysis of Industrial Cluster Learning. The learning system in the industrial cluster is shown in Figure 2.

TABLE 6: 2017-2019 per capita annual dairy consumption expenditure of rural households in a certain province with different income levels.

Income grouping	2017	2018	2019	Cumulative growth (%)
National average (yuan)	9.11	11.35	12.34	25.1
0–400	2.38	7.75	12.40	269.9
400–600	1.10	15.66	10.00	1250.8
600–800	4.00	3.57	4.40	4.6
800–1000	3.55	4.3 9	7.12	75.2
1000–1200	4.80	5.14	3.54	-29.9
1200–1500	7.76	7.41	6.12	-28.1
1500–1700	3.31	7.85	6.28	107.3
1700–2000	6.64	7.93	7.3 6	4.9
2000–2500	9.50	9.44	12.76	25.8
2500–3000	9.04	14.46	8.86	14.3
3000–3500	10.92	17.38	12.22	22.2
3500–4000	18.49	12.84	12.56	-38.9
4000–4500	16.57	17.51	15.72	-11.4
4500–5000	15.08	14.27	13.98	-14.3
>5000	20.79	21.3 7	27.02	20.8
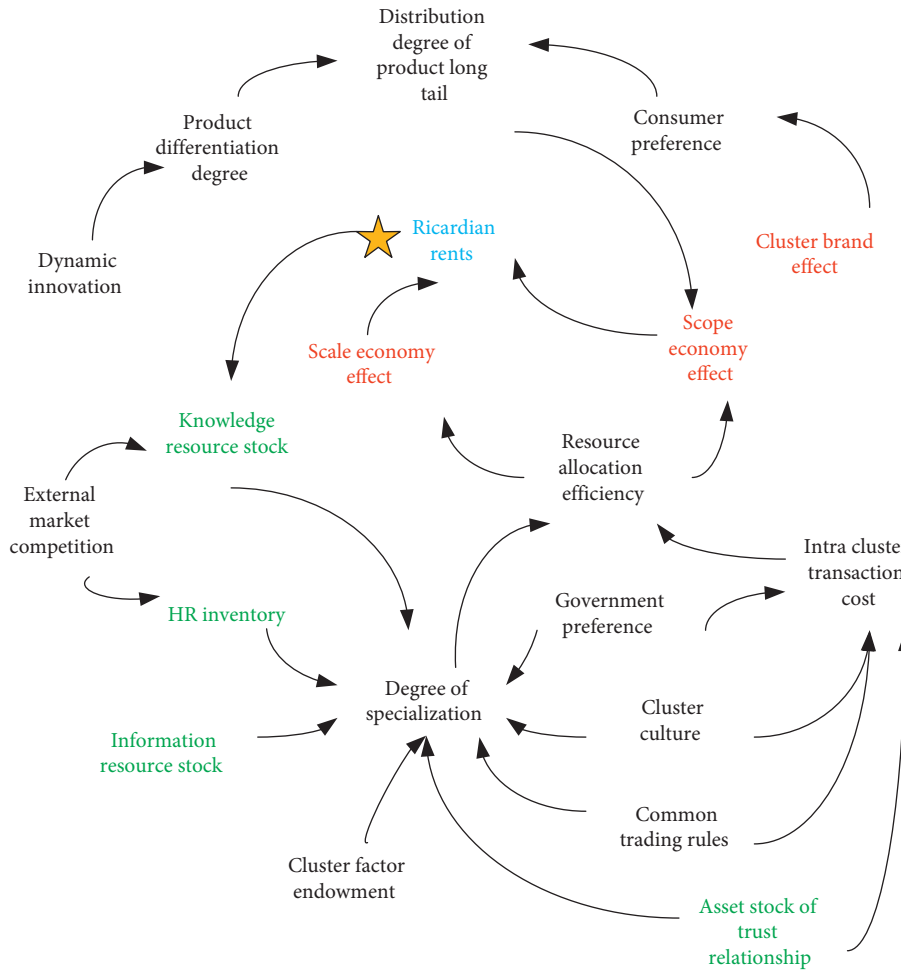
FIGURE 5: Ricardo's causality diagram of the rent subsystem.

TABLE 7: The main variables of Ricardo's rent system and their properties.

	Variable name	Nature	Initial value
	Ricardo rent	Li (t)	0
	Incremental heterogeneous resources	Rl (t)	0.0126
	External market competition	A (t)	0.5
Ricardo rent system	Stock of knowledge resources	A2 (t)	100
	Human resources stock	a1	100
	Information resource stock	a2	100
	Cluster factor endowment	a3	0.1
	Common trading rules	a4	1

3.4.2. Industrial Cluster Innovation and Power System. Innovation can bring development and competition for industrial clusters, and competitors can continuously improve the strength of the enterprise, so innovation is very important between enterprises. Many components can promote the innovation and development of an enterprise. We can roughly divide them into internal and external components. The internal and external components interact and influence each other to form a new driving force to promote the innovation of the cluster, as shown in Figure 3.

4. Dairy Industry Cluster Ecosystem Simulation Research Design

4.1. Analysis of Dairy Industry Development and Cluster Status

4.1.1. Overview of Price Competition among Dairy Companies. In recent years, the competition among the dairy industries in Western country is mainly reflected in prices. Many companies will reduce prices to increase their competitiveness. Some companies will choose to hold some promotional activities or give away some gifts to improve their

TABLE 8: Main variables and properties of the cluster ecological rent system.

	Variable name	Nature	Initial value
	Brand devaluation	R4 (t)	0.437
	Leading corporate brand investment	a24	3
	Proprietary resource share	a25	0.137
	Product novelty	A12 (t)	0.245
	Brand effect of leading cluster enterprises	A13 (t)	0.314
	Cluster brand effect	a26	0.56
Cluster ecological rent system	Cluster ecological rent	L5 (t)	0
	Competitive advantage	R5 (t)	0.2
	R&D subsidy policy	C1	
	Tax incentives	C2	
	Personnel policy	C3	
	Platform construction policy	C4	
	Brand building policy	C5	

TABLE 9: 2010-2019 China's dairy industry enterprise data.

Year	Total profit of private dairy industry enterprises (10,000 yuan)	Annual average number of employees in private dairy industry enterprises (10,000 people)	Per capita profit of private dairy industry enterprises (yuan)
2010	638200	61.04	10455.44
2011	970900	67.43	14398.64
2012	1399600	80.22	17447.02
2013	1575700	79.81	19743.14
2014	2454200	90.42	27142.23
2015	2691400	84.22	31956.78
2016	3686900	102.065	36123.06
2017	4032000	97.57	41324.18
2018	4362000	113.71	38360.74
2019	4332300	110.92	39057.88

TABLE 10: Validity test of industrial cluster ecosystem system dynamic model.

Year	Per capita profit of dairy cluster enterprises in a province (yuan)	Per capita profit of national dairy cluster enterprises (yuan)	Cluster excess profit (cluster ecological rent) (yuan)	System dynamic model simulation value	Error rate (%)
2010	14998.20	10455.44	333.01	364.44	9.44
2011	18359.25	14398.64	599.56	496.57	−17.18
2012	21803.10	10455.44	912.23	814.14	−10.75
2013	29336.00	14398.64	2059.96	2147.01	4.23
2014	14998.20	10455.44	2193.77	2419.81	10.30
2015	18359.25	14398.64	2204.41	2609.11	18.36
2016	40108.78	36123.06	3985.72	4354.62	9.26
2017	46906.44	41324.18	5582.26	4834.17	−13.40
2018	42296.40	38360.74	3935.66	4298.41	9.22
2019	43023.15 overall goodness-of-fit R^2	39057.88	3965.27	4414.71 91.94%	11.33

competitiveness. In the years of competition, the market value of different dairy industries and the prices of ordinary dairy products will also fluctuate greatly. See Table 1 for specific data.

See Table 2 for analysis of price changes in dairy products in different regions and different brands.

The data image corresponding to Table 2 is shown in Figure 4.

4.1.2. Questionnaire Survey-Specific Plan. Many people in China only tasted milk after the founding of New China.

Therefore, after the founding of New China, the Chinese country's dairy industry has developed rapidly. Among them, the per capita consumption of dairy products by urban residents in all provinces, cities, and towns across the country in 2013 and 2017 expenses is shown in Table 3.

The data in the table show that the consumption of dairy products in our country is at a relatively low level, not only much lower than some developed countries in Europe and North America, but also much lower than many Asian countries. The per capita consumption of liquid milk in major countries in the world is shown in Table 4.
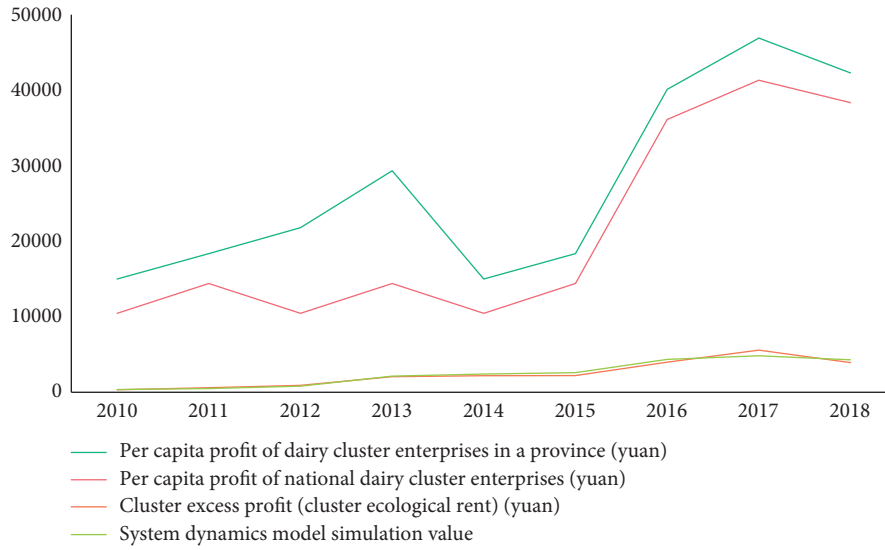
FIGURE 6: Model mean error analysis bubble chart.

TABLE 11: 2016-2019 average technical efficiency of dairy companies.

Index	2015	2016	2017	2018	2019	Mean
Technical efficiency	0.676 6	0.6814	0.7158	0.775 7	0.848 8	0.758 0
Pure technical efficiency	0.764 2	0.784 1	0.8461	0.869 3	0.900 2	0.848 9
Scale efficiency	0.885 4	0.869 0	0.8460	0.892 3	0.9429	0.891 6

TABLE 12: Distribution of technical efficiency of dairy companies in 2018.

Efficiency value distribution	Number of enterprises	Percentage
1	5	13.16
$0.9 \leq \theta < 1$	6	15.79
$0.8 \leq \theta < 0.9$	13	34.21
$0.7 \leq \theta < 0.8$	12	31.58
$0.6 \leq \theta < 0.7$	2	5.26
Total	38	100.00

TABLE 13: Distribution of pure technical efficiency of dairy companies in 2019.

Efficiency value distribution	Number of enterprises	Percentage
1	12	31.58
$0.9 \leq \sigma < 1$	7	18.42
$0.8 \leq \sigma < 0.9$	12	31.58
$0.7 \leq \sigma < 0.8$	7	18.42
Total	38	100.00

TABLE 14: Distribution of scale and efficiency of dairy companies in 2018.

Efficiency value distribution	Number of enterprises	Percentage
1	6	15.79
$0.9 \leq \omega < 1$	23	60.52
$0.8 \leq \omega < 0.9$	6	15.79
$0.7 \leq \omega < 0.8$	3	7.90
Total	38	100.00

TABLE 15: Distribution of changes in returns to scale of dairy companies in 2019.

Scale efficiency	Number of enterprises	Percentage
CRS	11	28.95
IRS	26	68.42
DRS	1	2.63
Total	38	100.00

4.1.3. Group Structure of Dairy Product Consumption.
During the research process, the staff sampled the consumption of dairy products by urban and rural residents in a certain province from 2015 to 2017 and studied the differences and mathematical relationships between the consumption of dairy products by residents of different incomes. The specific results are shown in Tables 5 and 6.

4.2. The Overall Structure of the Evolution Model of the Industrial Cluster Ecosystem

4.2.1. Ricardo's Rent Subsystem Feedback Loop.
Instead, the enterprise's resource stock, specialization degree of engineering, resource allocation efficiency, etc., are taken. As the research objects, a circular relationship diagram is established, the relationship between them is analyzed, and the analysis results are shown in Figure 5.

The efficiency of resource allocation of enterprises is mainly affected by the specialization of labor and internal exchanges. Staff draw the cause tree to explore the influencing factors of the formation of industrial clusters. Through the observation table, we can intuitively get the

TABLE 16: Technical efficiency values and rankings calculated by the super-SBM model in various provinces and municipalities across the country from 2015 to 2019.

Area	2015	2016	2017	2018	2019	2015–2019	
	Super-SBM	Super-SBM	Super-SBM	Super-SBM	Super-SBM	Average	Sort
Beijing	0.616 1	0.625 4	0.655 3	0.681 9	0.645 3	0.644 8	15
Tianjin	0.591 1	0.6823	0.596 6	0.633 3	0.6178	0.624 2	19
Hebei	1.1878	1.174 8	1.177 0	0.762 7	1.006 5	1.061 8	2
Shanxi	0.810 0	0.7747	0.601 1	0.698 3	0.575 0	0.6918	12
Inner Mongolia	1.049 3	1.026 7	1.001 4	1.024 4	0.760 0	0.972 3	4
Liaoning	0.546 8	0.552 7	0.653 0	0.737 1	0.930 9	0.684 1	13
Jilin	0.627 7	0.482 9	0.434 5	0.482 7	0.4212	0.489 8	30
Heilongjiang	0.8245	0.809 3	0.771 6	0.837 1	0.861 8	0.820 9	8
Shanghai	1.050 0	1.2228	1.1639	1.159 2	1.372 6	1.193 7	1
Jiangsu	0.698 0	0.658 1	0.599 6	0.618 9	0.610 1	0.637 0	17
Zhejiang	0.6243	0.6286	0.629 4	0.690 6	0.6271	0.640 0	16
Anhui	0.558 6	0.499 0	0.501 2	0.832 6	0.7747	0.633 2	18
Fujian	0.6535	0.552 3	0.521 5	0.640 6	0.602 0	0.594 0	23
Jiangxi	0.6148	0.5976	1.005 6	1.079 8	1.088 4	0.877 3	7
Shandong	0.832 1	0.840 9	0.856 5	1.084 9	1.008 4	0.924 6	5
Henan	0.497 7	0.658 4	0.741 5	0.878 6	0.825 1	0.720 3	10
Hubei	0.583 3	0.590 5	0.5264	0.625 8	0.5514	0.575 5	24
Hunan	1.000 7	0.9110	0.750 5	1.001 1	0.781 7	0.889 0	6
Guangdong	1.051 3	1.0221	1.015 2	1.023 8	1.033 6	1.029 2	3
Guangxi	0.570 2	0.561 2	0.5348	0.597 9	1.003 0	0.653 4	14
Hainan	0.6039	0.553 7	0.547 6	0.597 3	0.5277	0.566 1	26
Chongqing	0.641 8	0.556 0	0.537 4	0.608 2	0.644 8	0.597 6	22
Sichuan	0.611 7	0.626 0	0.640 4	0.487 0	0.648 5	0.6027	21
Guizhou	0.466 7	0.490 8	0.4128	0.419 1	0.412 3	0.440 3	31
Yunnan	0.543 7	0.515 1	0.511 2	0.638 6	0.6320	0.568 1	25
Tibet	0.313 1	1.060 4	1.110 4	0.513 0	0.485 6	0.6965	11
Shaanxi	0.640 0	0.7574	0.704 1	0.843 0	0.8173	0.752 4	9
Gansu	0.520 9	0.5484	0.498 9	0.586 2	0.463 7	0.523 6	28
Qinghai	0.520 8	0.438 9	0.466 8	0.726 4	0.557 3	0.542 0	27
Ningxia	0.6274	0.607 0	0.628 0	0.691 8	0.486 3	0.608 1	20
Xinjiang	0.533 6	0.4924	0.496 6	0.592 5	0.495 4	0.522 1	29
Average	0.6778	0.694 1	0.686 8	0.735 3	0.718 3	0.7025	

relationship between different influencing factors and influencing factors.

4.3. The Main Variables and Properties of the Evolution Model of the Industrial Cluster Ecosystem.

The rent system of this model and its cluster ecological rent system, the main variables, and types involved are shown in Table 7.

The staff will apply the parameter selection method in the mathematical research to the model to optimize the model, and the specific parameters will be determined after continuous experiments. During the experiment, the staff will make a correction based on the preliminary parameter values obtained. The model is adjusted to a certain extent. If the function of the model does not change significantly before and after the adjustment, then this parameter is the final parameter, as shown in Table 8.

5. Simulation Results and Analysis of the Dairy Industry Cluster Innovation Ecosystem

5.1. Fitting Test of the Dairy Industry Cluster Ecosystem Evolution Model.

The staff will adopt different testing methods to test whether the model is valid and whether the fit of the parameters meets the requirements:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y}_l)^2}{\sum_{i=1}^{n} (y_i - \overline{y}_l)^2}. \tag{26}$$

Different mathematical symbols in the formula represent different meanings.

In this study, industries in a certain area were randomly selected as objects to test the fit. The test process is shown in Table 9.

Table 10 shows that the selected model has practical meaning and is very accurate.

The values in Figure 6 represent the errors generated by the model, and these errors are within the tolerable range, so the model is credible and relatively accurate.

5.2. Model Data Extraction and Analysis.

Table 11 reflects the changes in the average technical efficiency of my country's dairy industry from 2016 to 2019. Through the table, it is not difficult to see that the per capita efficiency levels of different companies have changed significantly.

TABLE 17: Evolution stage characteristics of the industrial cluster ecosystem based on simulation.

Cluster development history	Segmentation basis	Characteristics of each course
The initial phase	None of the rents appeared	With Ricardo's rent as the mainstay, the advantages of clusters begin to manifest
Progression stage	A small amount of rent	The rents are evenly matched
Perfect stage	The relationship between rents begins to change	Ricardo's rent status has been replaced by the other three rents; the overall level of the cluster is the highest
Metamorphosis stage	Changes in different rent levels	Cluster level decline

Table 12 shows us the distribution of technical efficiency of various dairy companies in my country in 2018.

The staff summarized the approximate level of technical efficiency of my country's dairy companies in 2019, as shown in Table 13.

Based on the above data, the staff has integrated the distribution of the scale of my country's dairy companies in 2019. For detailed data, see Table 14.

The staff collected relevant data on the scale and marketing profit of each dairy industry in 2019 and integrated them into a table (Table 15).

Table 16 shows the calculation status and efficiency of the super-SBM model in each province across the country from 2014 to 2018.

It can be drawn from Table 16 the ranking of super-SBM in various provinces and municipalities across the country from 2014 to 2018.

5.3. Simulation Results and Analysis of the Dairy Industry Cluster Ecosystem Evolution Model. The characteristics of the evolution stage of the industrial cluster innovation ecosystem based on simulation are summarized in Table 17.

6. Conclusion

With the development of Internet technology, the Internet has played an increasingly important role in industrial clusters, not only affecting the development path of industrial clusters but also changing the development mode of industrial clusters. It can be said that the Internet has a historical role in industrial clusters. In the development process of industrial clusters, there are not only golden opportunities but also severe challenges. We should seize the opportunities without being afraid of the challenges. For industrial clusters, the development of network technology can promote changes in the industrial environment, thereby promoting the development of industrial clusters. In the development process of industrial clusters, the most important thing is practice. Industrial clusters can continuously improve and find and solve problems in practice. Through practice, the development of industrial clusters can be more mature, so that they can adapt to the new environment and welcome new ones. Challenge: the important role of industrial clusters in industrial development has also attracted the attention of many scholars. In recent years, more and more scholars have begun to take industrial clusters as their research direction.

This study mainly analyzes and studies the system and marketing profit level of my country's dairy industry in the form of a combination of research and theory and uses this as a sample to analyze the factors that will have a significant impact on industrial clusters in the process of entrepreneurial cooperation.

The preliminary conclusion drawn from the research is that factors such as government policies, industrial creativity, and industrial scale will all have a greater impact. To promote the long-term development of enterprise clusters, it is necessary to increase innovation and government support.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. H. Jia and J. P. Yang, "The review of leading firms' stimulative effect in the evolution of industrial cluster," *Review of Industrial Economics*, vol. 6, no. 1, pp. 129–136, 2007.

[2] J. Wei, *Industrial Cluster: Innovation Systems and Technological Learning*, Science Press, Beijing, China, 2003.

[3] Z. T. Wang, C. S. Zhang, and Z. W. Zhang, "Characteristics and optimization of collaborative innovation model for small and medium-sized technology-based enterprise—a case study of zhongguancun," *Academics*, vol. 8, pp. 239–244, 2015.

[4] M. M. Leng, *Competition and Cooperation in Supply Chains: Game-Theoretic Models*, McMaster University, Hamilton, Canada, 2005.

[5] R. Martin and P. Sunley, "Conceptualizing cluster evolution: beyond the life cycle model?" *Regional Studies*, vol. 45, no. 10, pp. 1299–1318, 2011.

[6] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[7] E. Alvarez-Garrido and G. Dushnitsky, "Are entrepreneurial venture's innovation rates sensitive to investor complementary assets? Comparing biotech ventures backed by corporate and independent VCs," *Strategic Management Journal*, vol. 37, no. 5, pp. 819–834, 2016.

[8] E. H. Hu, *Theoretical and Empirical Research on Innovation Behavior of Enterprise Clusters-Based on Complex Adaptive System Theory*, Science Press, Beijing, China, 2007.

[9] X. Zhang, Z. Zheng, K. Huang, and P. Wang, "Organizational culture, inter-organizational learning ability and innovation performance of the technology alliance of small and medium

enterprises," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management*, vol. 87, pp. 29–33, Bangkok, Thailand, December 2013.

[10] J. Li and J. Xing, "Why is collaborative agglomeration of innovation so important for improving regional innovation capabilities? A perspective based on collaborative agglomeration of industry-university-research institution," *Complexity*, vol. 2020, Article ID 7049606, 21 pages, 2020.

[11] T. Yıldız and Z. Aykanat, "Clustering and innovation concepts and innovative clusters: an Application on Technoparks in Turkey," *Procedia-Social and Behavioral Sciences*, vol. 195, pp. 1196–1205, 2015.

[12] M. Expósito-langa, J.-V. Tomás-Miquel, and F. X. Molina-Morales, "Innovation in clusters: exploration capacity, networking intensity and external resources," *Journal of Organizational Change Management*, vol. 28, no. 1, pp. 26–42, 2015.

[13] D. D. Fundeanu and C. S. Badele, "The impact of regional innovative clusters on competitiveness," *Procedia - Social and Behavioral Sciences*, vol. 124, pp. 405–414, 2014.

[14] L. Y. Ma and X. M. Li, "Does science and technology finance policies promote regional innovation? Quasi-natural experiment based on the pilot policy of combining science and technology with finance," *China Soft Science*, vol. 33, no. 12, pp. 30–42, 2019.

[15] J. H. Li, "Distribution and cultivation strategies of innovative industrial clusters in China," *Reforma*, vol. 3, pp. 98–110, 2020.

[16] Y. Zhou, Y. Guo, and Y. Liu, "High-level talent flow and its influence on regional unbalanced development in China," *Applied Geography*, vol. 91, no. 2, pp. 89–98, 2018.

[17] H. J. Sun, L. N. Zhang, and S. G. Wang, "Agglomeration of scientific and technological talents, spatial spillover and regional technological innovation: a partial differential method based on spatial Dubin model," *Science of Science and Management of Science and Technology*, vol. 40, no. 12, pp. 60–71, 2019.

[18] X. Y. Ji, "Does financial geography affect urban innovation capability?" *Industrial Economics Research*, vol. 1, pp. 114–127, 2020.

[19] R. Motzek, *Motivation in Open Innovation*, AV Akademikerverlag GmbH & Co. KG, Saarbrücken, Germany, 2012.

[20] E. Fakhrutdinova, S. Mokichev, and J. Kolesnikova, "The influence of cooperative connections on innovation activities of enterprises," *World Applied Sciences Journal*, vol. 27, no. 2, pp. 212–215, 2013.

[21] S. Liyanage, "Breeding innovation clusters through collaborative research networks," *Technovation*, vol. 15, no. 9, pp. 553–567, 1995.

[22] D. Yao and J. Whalley, "The China (Shanghai) pilot free trade zone: background, developments and preliminary assessment of initial impacts," *The World Economy*, vol. 39, no. 1, pp. 2–15, 2016.

[23] J. H. Guo, M. N. Guo, S. F. Guo, and S. S. Zhang, "Dose the innovation policy pilot promotes enterprise innovation effectively in China? Empirical evidence from the construction of national independent innovation demonstration zones," *Industrial Economics Research*, vol. 2, pp. 56–70, 2021.

[24] B. Kaviarasan, R. Sakthivel, and S. Abbas, "Robust consensus of nonlinear multi-agent systems via reliable control with probabilistic time delay," *Complexity*, vol. 21, no. S2, pp. 138–150, 2016.

Hindawi

Research Article

Research on Multiplayer Posture Estimation Technology of Sports Competition Video Based on Graph Neural Network Algorithm

Xiaoping Guo (ID)

Shaanxi Normal University, Xi'an, Shaanxi 710000, China

Correspondence should be addressed to Xiaoping Guo; gxp2006110@snnu.edu.cn

With the explosive growth of the number of sports videos, the traditional sports video analysis method based on manual annotation has been difficult to meet the growing demand because of its high cost and many limitations. The traditional model is usually based on the target detection algorithm of manual features, and the detection of human posture features is not accurate. Compared with global image features such as line features, texture features and structure features, local image features have the characteristics of rich quantity in the image, low correlation between features, and will not affect the detection and matching of other features due to the disappearance of some features in the case of occlusion. Referring to the practice of Deep-ID network considering both local and global features, this paper adjusts the traditional neural network, and combines the improved neural network with the human joint model to form a human pose detection method based on graph neural network, and then applies the algorithm to multiperson human pose estimation. The results of several groups of comparative experiments show that the algorithm can better estimate the human posture in sports competition video, and has a good performance in solving multiperson pose estimation in sports game video.

1. Introduction

With the development of society, more and more sports videos have entered people's daily life. Analyzing these sports videos can not only bring more wonderful content to the audience but also find out the shortcomings of athletes for improvement. Accurate athlete detection and pose estimation are important links in sports video analysis. The existing human target detection and pose estimation algorithms have achieved good performance in general human detection tasks, but they will detect athletes and spectators at the same time in sports video, so they cannot further distinguish athletes' targets, which will interfere with the subsequent video analysis. At the same time, the sports video data with human body annotation is scarce, and the cost of obtaining a model suitable for the field of sports video is high. Athletes are a special case of general human body detection task. If the general human body detection model can be used to detect and estimate athletes in sports video, it can undoubtedly save a lot of cost. The analysis and

understanding of sports video have a huge market demand [1], and can provide standard teaching cases for sports lovers. In addition, the statistics of various data of athletes in the video can not only help athletes improve their technical level, but also adjust the tactical deployment for the whole team, and help coaches and athletes improve the strength of the team to a certain extent. The demand for sports video analysis and understanding is increasing.

The existing target detection and human pose estimation technologies have achieved good performance in the general scene detection task based on pictures, but there are few algorithms and data for target detection in sports video scenes. The usual approach is to label these data, and then train the target detection algorithm with new data to obtain the detection model [2]. Although this method can achieve results, it needs to retrain the model every time, which requires a lot of labeling cost and training cost. There are many kinds of sports videos, and a video contains a large number of pictures [3], so it is undoubtedly expensive to label the video. When analyzing sports video, the focus is on

athletes. The target detector trained based on general target detection task itself has the ability to recognize people. Similarly, human posture estimation for sports video is also one of many human posture estimation scenes. The human posture detector based on general scene training also has the ability to recognize human posture in sports video. Ifmmc1 the trained general detector can be directly used for the detection of athletes, it can undoubtedly greatly reduce the cost [4]. However, in sports video, people are crowded, occlusion often occurs between people, motion blur is serious, and the general target detector does not have the ability to distinguish ordinary spectators and athletes [5], which brings challenges to the use of general detector in sports video scenes.

This paper mainly studies the athlete detection and human posture estimation in sports video. Based on graph neural network algorithm and human posture estimation algorithm, combined with the characteristics of sports video, the target detection and human posture estimation model based on general data set training is transferred to the field of sports video, The purpose is to reduce the training and labeling cost of athlete detection and pose estimation for sports video, and improve the performance of athlete detection and human pose estimation in sports video. In this paper, the task of athlete detection and human posture estimation for sports video mainly focuses on the problem of multiperson motion in video. The LSP data set with 14 labeled joint points is adopted to meet the characteristics of multiperson motion and recognizable posture. The performance of this algorithm is compared with four mainstream methods of human pose estimation in multiple databases to verify the effectiveness of this algorithm in the field of multiperson pose estimation.

2. Related Work

Target detection is a basic task in the field of computer vision, and has always been a hot research topic. Before the introduction of depth neural network, the target detection algorithm is mainly realized by extracting the manual features of the image for feature matching [6]. Many image manual features such as sift and surf have achieved good results. A lot of work has tried to design different manual features for target detection according to the characteristics of detection targets, such as V-J detection, hog detection [7], and DPM [8] algorithm. The target detection algorithm based on manual features mainly uses the information of target color, texture and edge structure to recognize objects. It can achieve good results in a relatively simple background. However, in practice, the background changes in different fields are complex, and the image is often affected by illumination and resolution, resulting in target deformation and blur, so it is difficult to cover all the cases by manual features. The generalization ability of target detection algorithm based on manual features is weak. Therefore, the performance in complex scenes is not very good.

With the development of deep learning, human pose estimation algorithm based on deep learning has become the mainstream. The human posture estimation algorithm based on deep learning regards human posture detection as a key point regression problem, trains through a large number of data with joint point category and position annotation, and finally obtains a model that can predict the position and category of human joint points [9]. In the early days, people mainly studied single person pose estimation, that is, there is only one human body in the detected picture. The single person Newell et al. [10] designed an hourglass like network structure for attitude detection, and added a supervision signal in the middle of the network. The single person pose task only detects one human pose, which has simple picture background and less interference. The existing single person pose estimation algorithms have achieved good performance and achieved an accuracy of more than 93% on the single person pose estimation data set MPII. However, in practice, most pictures have multiple human bodies, so the single person pose estimation algorithm is no longer applicable. Compared with single person pictures, multiperson pictures have complex background and serious interference between people. The attitude estimation algorithm for multiperson faces more problems [11]. The existing multiperson attitude estimation algorithms are mainly divided into two categories: top-down multiperson attitude estimation algorithm and bottom-up multiperson attitude estimation algorithm. The top-down multiperson attitude estimation algorithm first detects the human body in the figure through target detection, then cuts out a single human body according to the detection frame and sends it to the single person attitude detector to obtain the human body attitude, and finally puts the single person's attitude back to the corresponding position in the original figure to obtain the human body attitude of all people [12]. Alphapose [13] algorithm introduces the spatial transformer network module to optimize the positioning accuracy of the detection frame and achieve a more accurate multiperson attitude estimation effect. Chen et al. [14] designed a cascaded pyramid network to complete multiperson attitude estimation through multiple regression and adjustment. The top-down multiperson attitude estimation algorithm decomposes the multiperson attitude detection problem into multiple single person attitude detection problems, which is realized by target detection and single person attitude detection. This method is simple and effective, but its performance will be affected by the target detection results. Because of the limitations of the target detection frame, when there are multiple people gathering and occlusion, there will be multiple human bodies in the target detection frame. At this time, the effect of the top-down multiperson pose estimation algorithm will be affected.

To sum up, the existing target detection algorithms are mainly based on convolutional neural network [15], and have achieved good performance on the general target detection data set. At the same time, the algorithm needs a lot of data-driven, and the cost of training a new model is high [16]. When target detection is carried out in some specific fields, the existing methods usually design and retrain the target detection model based on the data in this field. Although it can achieve good performance, it costs a lot. However, in practice, the specific target to be detected is

often a specific example of a certain class in the general target detection task, such as van detection and athlete detection, which are special cases of vehicle detection and human body detection in the general target detection task. At this time, if the trained general target detection model can be used to detect these specific targets, it can not only reduce the training cost, but also improve the application scope of the general detection model. The existing algorithms pay more attention to how to improve the detection performance in specific fields, and there is less research on how to migrate the general target detection model to specific fields to achieve specific targets.

3. Multiperson Pose Estimation Algorithm Based on Graph Neural Network

3.1. Graph Neural Network Algorithm.
The structure of neural network mainly includes convolution layer and pooling layer. The function of convolution layer is to extract abstract features by convolution kernel and reduce the number of parameters by local perception and weight sharing. The function of the pool layer is to aggregate and count the feature map extracted from the convolution layer, and further reduce the number of neurons by down sampling [17]. When using convolution check sample image block to extract abstract features, different regions of the sample image are treated equally, but in fact, the joint is only located in the central region of the positive sample image, so the central region plays a greater role in identifying the joint than other regions. Therefore, when extracting features, we should pay more attention to the central region and weaken the edge region. Therefore, this paper improves the convolution operation of neural network, and assigns different weights to the convolution operation in different regions of the image. Due to the similarity of the same species, although the appearance features of the same joint image block of different human bodies will be different due to different body shapes and postures, the difference is not very large, especially in the central area of the image, so they have similar local features. Since the same joint image blocks of different human bodies have similar local features, their global features will also have great similarity. The global features can also play a certain role in identifying joints. Therefore, in identifying joints, we should consider not only the local features of the image, but also the global features. Therefore, this paper improves the structure of the traditional neural network by referring to the deep id network in the face recognition algorithm, which considers both local and global features, and the global and local features of the image were extracted and combined as image features. The image features include 32-dimensional local features and 1-dimensional global features, which are mainly determined by local features, as shown in Figure 1.

Firstly, the local response of the color image is normalized to improve the generalization ability of the neural network. According to the size of the input vector and the global features of the output one-dimensional image, the neural network structure set in this paper is shown in Table 1, including 3 convolution layers, 4 pooling layers, 1 lead-

in layer and 3 full connection layers. Some of the parameters are shown in Table 1.

The improved graph neural network contains three convolution layers. The meaning of the parameters of each convolution layer in Table 1 is similar. For example, the output parameter of convolution layer 3 is $9 \times 9 \times 32$, where 32 is the number of convolution kernels. A convolution kernel corresponds to a feature extraction method. After convolution, a feature map is obtained, 9×9 represents the size of the characteristic image obtained after convolution, and the convolution step of all convolution layers is 1×1. The convolution layer activation function adopts the modified linear unit (ReLU) with strong anti-overfitting ability. ReLU will make the output of some neurons 0, which leads to the sparsity of the network, reduces the interdependence of parameters and alleviates the over fitting problem.

The learning algorithm of the improved CNN adopts the random gradient descent algorithm, and the objective function is set as

$$O(W) = \sum_{i=1}^{N} \left(f^i(W) - d^i \right), \tag{1}$$

where N is the number of samples, $f^i(W)$ is CNN output, d^i is the sample classification label, and positive and negative samples are 1 and 0 respectively. The improved graph neural network contains three full connection layers, and the neurons in the layer are connected with all the neurons in the upper layer. In front of the full connection layer is the introduction layer, but the convolution layer 3 and the pooling layer 4 are introduced into the full connection layer. The activation function of the first two fully connected layers adopts the modified linear unit (ReLU), and the fully connected layer 3 adopts the logistic regression function, and the output is the joint label.

3.2. Human Pose Estimation of Single Frame Image.
In this section, the video is divided into a series of single frame images. For each frame image, the position of each joint point of the upper body of the human body is obtained, and the posterior edge probability distribution of the wrist is calculated. Maximum likelihood estimation can be understood as a map estimator when the prior probability is uniformly distributed, and obtains the point estimation of the difficult to observe quantity according to the empirical data. The maximum a posteriori estimation is integrated into the prior distribution of the quantity to be estimated, which can be regarded as the regularized maximum likelihood estimation. In order to realize the human body pose estimation of a single frame image, the most commonly used graph structure model is used to represent the basic human body architecture, and the reasoning algorithm is used to obtain the exact joint position by transmitting messages between the parent node and the child joint point.

Figure 2 shows a schematic diagram of a structural model. The structure of human body is described by undirected graph G(E, V), where the vertex $v_i \in V$ represents a
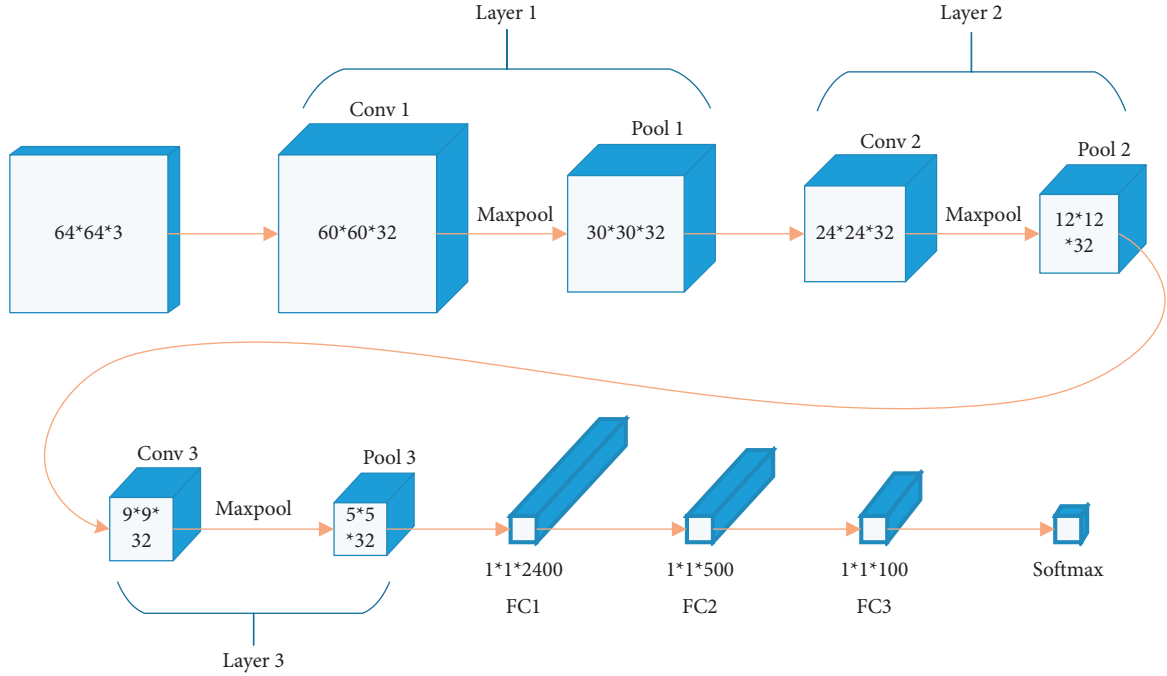
Figure 1: Improved graph neural network structure.

Table 1: Figure structural parameters of neural network.

Layer	Conv1	Pool1	Conv2	Pool2	Conv3
Input	$64 * 64 * 3$	$60 * 60 * 32$	$30 * 30 * 32$	$24 * 24 * 32$	$12 * 12 * 32$
Output	$60 * 60 * 32$	$30 * 30 * 32$	$24 * 24 * 32$	$12 * 12 * 32$	$9 * 9 * 32$
Size	$5 * 5 * 3$	$2 * 2 * 1$	$5 * 5 * 32$	$2 * 2 * 1$	$5 * 5 * 32$
Layer	Pool3	Pool4	FC1	FC2	FC3
Input	$9 * 9 * 32$	$5 * 5 * 32$	$1 * 1 * 2400$	$1 * 1 * 2400$	$1 * 1 * 500$
Output	$5 * 5 * 32$	$1 * 1 * 32$	$1 * 1 * 2400$	$1 * 1 * 500$	$1 * 1 * 100$
Size	$2 * 2 * 1$	$5 * 5 * 1$	—	—	—
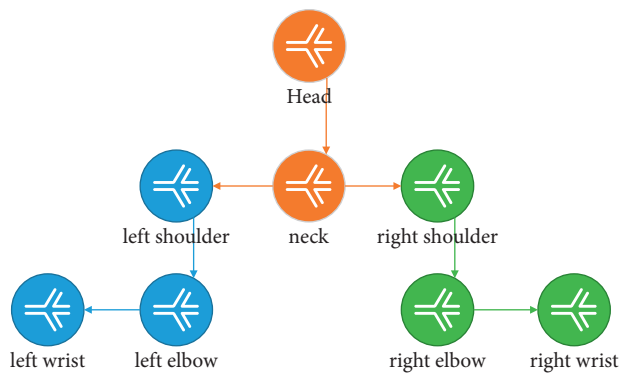
Figure 2: Structural model of human upper body diagram.

human body part, and the edge line e represents the adjacent joint point vi and vj For example, in the human body structure, there is a physical connection between the left wrist and the left elbow. Use the human body configuration $P = \{p^1, p^2, \ldots p^i, p^{n-1}, p^n\}$ to describe the posture of the human body, where $p^i = (x^i, y^i, \theta^i)$ is used to describe the human body part vi. (xi, yi) represents the coordinates of the center point of the rectangular box of part v^i, θ^i Indicates component v_i The direction of the rectangle. Given a frame of image I, the human body configuration P is defined as

$$P = k \cdot e^{\sum \varphi(p^i, p^j) + \sum \phi(p^i, I)}, \quad (i, j) \in E, i \in V, \quad (2)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ are potential energy functions. Paired term potential energy function $\varphi(p^i, p^j)$ model the geometric constraint relationship between two adjacent joint points by encoding the human body a priori model. $\Phi(p^i, I)$ indicates the feasibility that the component v_i is located at the p^i position in the image I.

The best posture of human body in a single frame image is obtained by formula (1). Generally, the graph structure model has a tree structure, so two different reasoning algorithms can be used in reasoning. When reasoning with the maximum a posteriori (map), select the maximum product method to obtain the globally optimal maximum a posteriori estimation and the sum product method to obtain the edge distribution of the wrist, and take the edge distribution of the wrist as the a posteriori distribution probability of the tracking algorithm in the next stage.

In order to process video sequences more effectively, each frame needs to be read. Processing video frames mainly applies some processing functions to each video frame. As shown in Figure 3, the data processing module first divides the input sports video into frames, and unifies the size of the video frame to 1000×600 pixels, then detect people in each frame of the video based on the general target detector, then input the detection results obtained by the general target detector into the trained example feature metric classifier to obtain the athlete's detection results, and then fuse the multi frame detection results based on the video time domain context to obtain the optimized athlete's detection results.

3.3. Graph Neural Network Structure Model for Multiperson Attitude Estimation.

In order to ensure the recall rate of athletes, this paper takes all the detection frames of the general target detector as candidate frames when the threshold is 0. Extracting the features of the detection frame is to extract the features of the image content contained in the detection frame. The CNN network in the detection model has extracted the features of the whole image. According to the position of the detection frame, the ROIAlign method is used to extract the features of the detection frame.

Firstly, the feature map is resized to the size of the original map by bilinear interpolation, and then crop is performed on the feature map according to the position of the detection frame. Then, the detection frame feature map max pool to a unified size. Finally, the detection frame feature map is linearly expanded to obtain one-dimensional detection frame features. The size of the feature map of the detection frame after Max pooling will change according to different convolution neural networks, so that the feature length of the one-dimensional detection frame is between 1024 and 2048. In particular, a two-stage target detection network with RPN structure, such as fast RCNN, will extract features from each detection frame for classification and frame regression, that is, each detection frame obtained by fast RCNN target detection algorithm has a unique depth feature corresponding to it. RPN network compares the predicted results with the real results, reflects the gap between the predicted and the real value through the user-defined "distance," and optimizes the "distance" through the optimization algorithm, so the prediction is closer to the real value. In this paper, these features are directly used to identify the candidate frame as the feature of the candidate frame, The frame feature dimension extracted by the fast RCNN model used in this paper is 2048 dimensions.

The multilevel feature diagram is shown in Figure 4. To identify athletes, that is to distinguish athletes and nonathletes, the essence is to classify them according to the characteristics of candidate boxes. A feature vector q is defined to represent athletes, so that the candidate frame features of athletes are close to Q and the candidate frame features of nonathletes are far from Q. in this way, a threshold can be set to distinguish athletes from nonathletes.

From the candidate frame selection and feature extraction module, the 2048 dimensional depth feature corresponding to each detection frame can be obtained. This feature can be used to distinguish different detection frames, but cannot distinguish athletes from nonmobilization. Therefore, it is necessary to carry out a linear transformation on these features. This linear transformation module transforms the 2048 dimensional candidate frame feature into a 512 dimensional new feature. Initialize a 512 dimensional feature vector Q to represent the athlete. By continuously optimizing the parameters of this linear transformation module, the similarity between the new features of the athlete candidate box and Q becomes larger, and the similarity between the new features of the nonathlete candidate box and Q becomes smaller. The calculation formula of the linear transformation module is: $R = \tanh(W f)$.

where f represents the original 2048 dimensional feature of the candidate box, R represents the 512 dimensional new feature after linear transformation, tanh is one of hyperbolic functions, and tanh () is hyperbolic tangent. In mathematics, the hyperbolic tangent "tanh" is derived from the hyperbolic sine and hyperbolic cosine of the basic hyperbolic function, and W is the weight parameter of the linear transformation module.

Firstly, initialize a 512 dimensional feature vector Q to represent the athlete, then select the candidate box and extract the features of the picture to obtain n 2048 dimensional candidate Box feature vectors F, and then linearly transform the candidate box features to have the same dimension as the athlete features to obtain n 512 dimensional candidate box features R. Then calculate the cosine similarity between each candidate frame feature R and athlete feature Q to obtain AI, where I represents the number of candidate frames, and the values are 1 to n. All the detection frames in a picture are regarded as a package, and the maximum cosine similarity between the candidate frame feature and the athlete feature is taken as the similarity S between the package and the athlete feature. Suppose Q represents an athlete, R represents a picture containing an athlete, then the positive packet similarity can be expressed as $s(Q, R)$, P represents a picture without an athlete, and the negative packet similarity can be expressed as $s(Q, P)$. Theoretically, the similarity between a picture containing an athlete and an athlete category is greater than a picture without an athlete, that is, $s(Q, R) > s(Q, P)$.
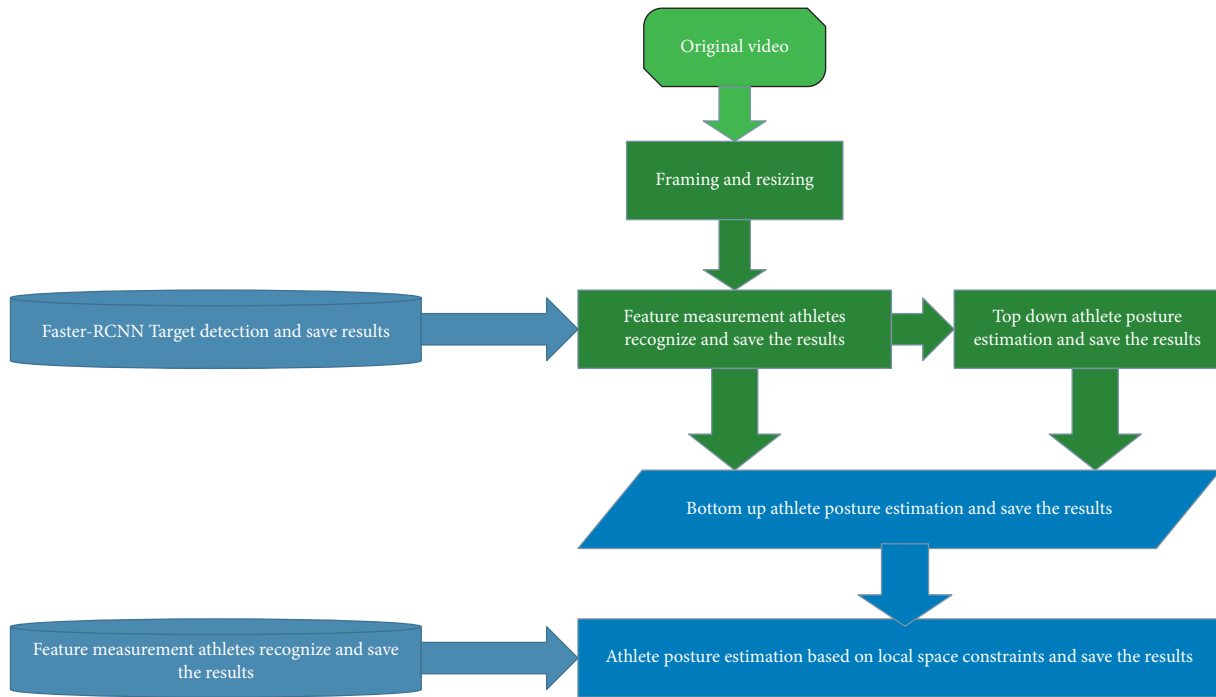
FIGURE 3: Data processing module.
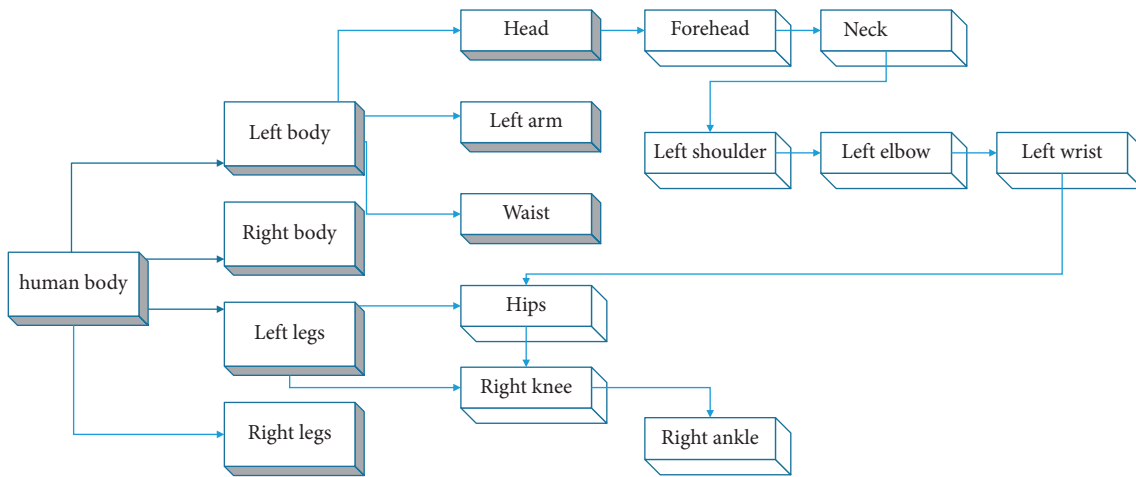
FIGURE 4: Multilevel characteristic diagram of athletes in sports competition.

The graph neural network is mainly composed of feature extraction layer, region recommendation layer, region of interest pooling layer and classification layer. Feature extraction layer and region recommendation layer have an important impact on the realization of network functions. Figure 5 is the network structure diagram of the neural network, which integrates the functions of feature extraction, candidate box selection, location refinement and classification into the same network. The designed structure corresponds to the feature extraction layer, region recommendation layer, region of interest pooling layer and classification layer respectively.

Given a natural image containing human posture, its size $P * Q$ needs to be reduced to a fixed size $M * N$ and then input into the graph neural network. $M * N$ images first pass through the feature extraction layer in the network, use this layer to extract the feature map describing human posture, and share the feature map describing human posture image in the subsequent region suggestion layer and region of interest pooling layer. The part of the graph neural network that is only directly connected with the feature extraction layer is called the region recommendation layer, which can be used to generate candidate frames of human parts. The region suggestion layer first generates rectangular boxes of different sizes through the convolution layer, then uses the classification function to judge whether the image block contained in the generated rectangular box is a human body part or a background pixel, and simultaneously calculates
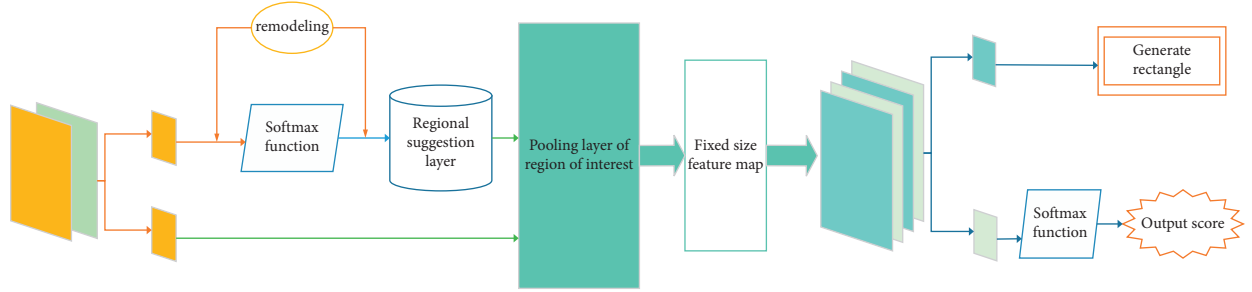
Figure 5: Neural network structure diagram of human posture in sports video.

the offset of the generated rectangular box to fine tune the position of the generated rectangular box. So far, the graph neural network has completed the positioning function of human parts. Considering that the region suggestion layer generates a rectangular box with non-fixed size, which cannot be directly used as the input of the classification layer, the region of interest pooling layer is introduced into the graph neural network. The region of interest pooling layer takes the original feature map and the generated rectangular box as the input at the same time, and extracts the feature vector of the image block in the generated rectangular box by blocking in the original human posture feature map, so as to ensure the output of the feature vector of fixed size. Finally, the fixed size feature vector is input to the classification layer, and the specific human body part type of the image block contained in the generated rectangular box is output.

Supervised learning is a machine learning task to infer a function from the marked training data. The training data includes a set of training examples. In supervised learning, Each instance consists of an input object (usually a vector) and a desired output value (also known as supervised signal) composition. Supervised learning algorithm is the function of analyzing the training data and generating an inference, which can be used to map new instances. An optimal scheme will allow the algorithm to correctly determine the class labels of those invisible instances. This requires that the learning algorithm is in a "reasonable." The way is formed from a training data to an invisible situation. Before training the network, the position of each part of the human body is marked in a rectangular box to realize the supervised learning of network parameters. When testing the network, the static image containing human body is directly input. After the graph neural network, the image with a series of rectangular detection boxes and the fractional probability vector corresponding to the rectangular detection box are output.

3.4. Evaluating Indicator. The athlete detection and evaluation is the same as the evaluation standard of general target detection. This paper evaluates by calculating the joint sum intersection ratio IoU (interaction over union) between the detection result frame and the marked frame. When IoU >0.5, it is considered that the athlete is correctly detected.

For the estimation and evaluation of athletes' posture, there are two evaluation methods of human joint points: PCK [18] (percentage of correct keypoints) and OKS [19]

(object keypoint similarity). This paper selects the commonly used OKS method, which will calculate the similarity between the two postures and give a score between 0 and 1. In this paper, the OKS scores of predicted posture and real posture are calculated. The OKs score is calculated as follows:

$$\text{OKS}_p = \frac{\sum_i \exp - d_{pi}^2 / 2S_p^2 \sigma_i^2 \delta(v_{pi} = 1)}{\sum_i \delta(v_{pi} = 1)}. \tag{3}$$

p represents the ID of the person in the growth truth, and i represents the ID of the keypoint. D_{pi} represents the Euclidean distance between each person in the growth truth and the predicted key point of each person. S_p represents the scale factor of the current person, which is equal to the square root of the area occupied by the person in the ground truth. σ_i represents the normalization factor of the i-th bone point. Therefore, this is obtained by calculating the standard deviation of all ground truths in the data set, reflecting the standard deviation when labeling the current bone point. v_{pi} represents whether the i-th key point of the p-th person is visible. δ function for selecting visible points for calculation. When OKS >0.5, it is considered that the athlete's posture is correctly detected. Athlete detection and posture estimation are both kinds of detection tasks. According to the evaluation method of detection tasks, this paper selects three common evaluation indexes, namely maximum recall rate, AP value and P-R curve, which are defined as follows. For the test results, there may be four situations between the predicted value and the actual value: TP, FP, FN, TN. TP and FP respectively indicate that the predicted positive sample and negative sample are true, FN and TN respectively indicate that the predicted positive sample and negative sample are false. Where P in the P-R curve represents precision and R represents recall. The formula are as follows:

$$P = TP(TP + FP)^{-1},$$
$$R = TP(TP + FN)^{-1}. \tag{4}$$

By setting different detection confidence thresholds, different detection results will be generated. The accuracy and recall rate of the detection results of each threshold can be calculated. The recall rate obtained by setting the confidence threshold to the minimum is the maximum recall rate. Calculate the accuracy and recall for each confidence threshold, and then take the recall as the abscissa and the

accuracy as the ordinate to draw the P-R curve, which can well reflect the performance change of the detector. The AP value of target detection can be obtained by calculating the integral of the curve in the interval [0, 1]. The AP value is an important index to measure the performance of target detection. The larger the AP value, the better the performance of the detector.

This article selects two objective evaluation criteria to objectively evaluate the experimental results of this algorithm, one is the percentage of detected joints (PDJ) and the other is the percentage of correct parts (PCP). The specific meaning of PDJ evaluation standard is that for each joint point, when the distance between the predicted position and the real position (calculated by Euclidean distance here) is less than the given threshold, the joint point is correctly located. The PDJ value of each joint point changes with the change of threshold value. The resulting curve depicts the change trend of joint point positioning accuracy, which is called PDJ curve. The specific meaning of PCP evaluation standard is that when the two joint points corresponding to both ends of a human part are correctly positioned, the part is correctly positioned.

4. Experiment and Analysis

4.1. Experimental Setup and Data Set. Based on the video time domain context, the effective fusion of the detection results of different video frames can effectively improve the recall rate of athlete detection, and solve the problem of missing detection caused by motion blur to a certain extent,

but it may also bring the risk of false detection. The optimization strategy of detection results based on time-domain context involves multiple super parameters. In order to ensure the optimization effect, these super parameters need to be set reasonably. The super parameters and value ranges involved are as follows: the value range of adjacent frame interval n is [0, 30], the value range of interframe similarity threshold s is [0, 64], the value range of local similarity threshold D is [0, 64], and the value range of supplementary detection frame confidence threshold T is [−1, 1]. In this paper, the grid parameter search method is used to experiment on NBA basketball video data set to determine the best parameter combination. In order to improve the efficiency of grid parameter search, the computational cost of parameter search can be reduced by removing some obviously unreasonable parameter ranges. According to experience, the search interval and search step of each super parameter are as follows: the search range of adjacent frame interval n is [0, 10], the search step is 1, the search range of interframe similarity threshold s is [0, 10], the search step is 1, the search range of local similarity threshold D is [0, 10], the search step is 1, the search range of supplementary detection frame confidence threshold T is [0, 1], and the search step is 0.1. After continuous parameter search, when the super parameters $N = 3$, $s = 5$, $d = 2$, $t = 0.3$, the sports competition video data set constructed in this paper can achieve good results. Parameter W represents the weight vector in the multilevel graph structure model. The process is as follows:

$$\min_{w}\langle\omega_{ij}, w_{ij}\rangle + \sum C \max\left[0, 1 - y_n\langle w, \Phi(I^n, l^n, t^n)\rangle\right], \quad y_n \in \{0, 1\}. \tag{5}$$

In the above formula, $\Phi(n)$ is the sparse feature of the n-th training image.

In this paper, the grid parameter search method is used to experiment on the data set to determine the best parameter combination. Grid search method is an exhaustive search method to specify parameter values. The optimal learning algorithm is obtained by optimizing the parameters of the estimation function through cross validation. That is, the possible values of various parameters are arranged and combined, and all possible combination results are listed to generate a "grid." Then each combination is used for SVM training, and the performance is evaluated by cross validation. After the fitting function tries all the parameter combinations, it returns an appropriate classifier and automatically adjusts to the best parameter combination. In order to improve the efficiency of grid parameter search, the computational cost of parameter search can be reduced by removing some obviously unreasonable parameter ranges. According to experience, the search interval and search step of each super parameter are as follows: the search range of adjacent frame interval n is [0, 10], the search step is 1, the search range of inter frame similarity threshold s is [0, 10], the search step is 1, the search range of local similarity

threshold D is [0, 10], the search step is 1, the search range of supplementary detection frame confidence threshold T is [0, 1], and the search step is 0.1.

At present, the mainstream human posture data sets used in the research include 2D and 3D. The 2D human posture data sets are shown in Table 2, BBG, and common data sets include MSCOCO, MPII, LSP, FLIC, posetrack and AI challenger.

Leeds sports pose is a sports pose data set, which is divided into sports, badminton, baseball, gymnastics, parkour, football, volleyball and tennis. It contains about 2000 pose notes, and the images are from the athletes of Flickr. Each image is a 3-channel color image, and each image is marked with 14 joint positions, and the left and right joints are always marked with human center. The LSP data set is collected from different motion scenes. Its extended version includes 11000 training samples and 1000 test samples. There are 14 labeled joint point types in the LSP data set, and the LSP data set does not need to be cropped, because the author of the data set provides preprocessed images. In the experiment, only the provided images need to be scaled to a fixed size. Therefore, the human posture estimation data set LSP is selected in this experiment, and other data sets such as

TABLE 2: Human pose estimate datasets.

Dataset	Feature
LSP	The data comes from sports category labels, image zoom, only annotated in each picture-one person, 14 key points
FLIC	Data comes from hollywood movies, 10 key points
MPII human pose	The data comes from YouTube videos, which covers 410 human activities, each image has an activity tag, and 16 key points
MSCOCO	The data comes from the internet, it contains various activities, 17 key points
Al challenge	The data is captured from the internet. It is currently the largest human pose image data set, and 14 key point
PoseTrack	The data comes from the MPII human pose data set and 15 key points
Buffy	The data comes from a TV show. Line segments are provided to indicate the location, and the size and direction of the body parts are also provided

FLIC, MSCOCO and Buffy pose are also used for comparison.

4.2. Evaluation of Multiperson Attitude Estimation Model. Figure 6 shows the loss curve during the training of the three-dimensional attitude estimation network. The training process is 300 rounds in total. It can be seen that the loss function (loss) value shows a decreasing trend with the increase of the number of training rounds, and tends to be stable after 250 rounds, indicating that the model training is gradually convergent.

We use a graph neural network algorithm for multiperson human pose estimation. The human posture estimation process is divided into two stages: detection component and positioning component [20]. In the first stage, the specific deep learning network is used as the human part detector to realize the detection function of human parts; In the second stage, a new spatial constraint relationship of human parts is designed to uniquely determine the position of each part from the candidate box of human parts, realize the positioning function of joint points, and finally complete the estimation process of human posture. In order to objectively verify the role of the component detection model based on graph neural network and the designed human component spatial constraint model in the whole algorithm, experiments on controlling a single variable are carried out on FLIC and Buffy databases, which can represent the overall performance of the traditional attitude estimation method. Green curve is the average PDJ curve obtained by hiding the spatial constraint model of human parts and directly taking the rectangular box position corresponding to the highest score probability vector output by fast R-CNN as the positioning result of human parts; The red curve is the average PDJ curve obtained by the algorithm. Comparing the green curve and purple curve longitudinally, it can be concluded that compared with the traditional methods, only using deep learning network to solve the problem of human posture estimation from the perspective of target detection has great advantages in positioning accuracy. In addition, as can be seen from Figure 7, both the component detection model based on fast R-CNN and the proposed human component spatial constraint model contribute to the final realization of accurate human pose estimation results.
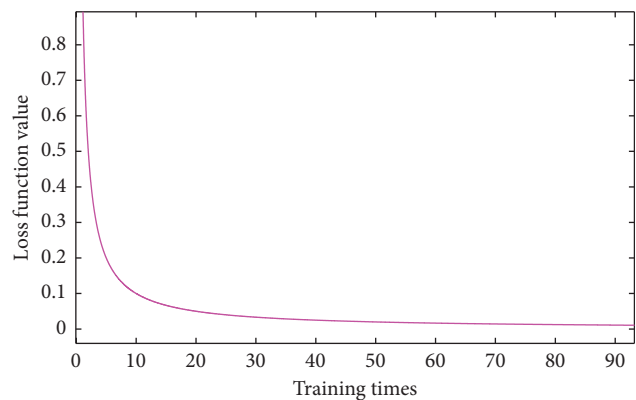
FIGURE 6: Loss curve of human posture estimation network.
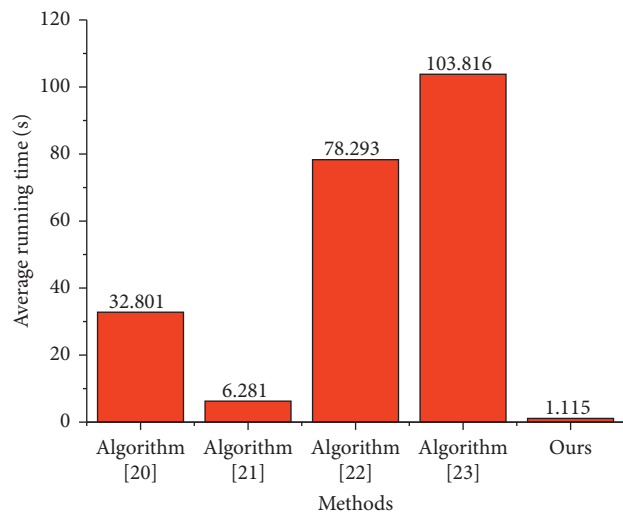
FIGURE 7: Running time comparison of human pose estimate.

In order to test the running time of this algorithm in the process of 3D human body pose estimation, this algorithm is compared with literature [20] and literature [21] under the same software and hardware conditions. The results are shown in Figure 7. The average time per frame of algorithm [20] is 32.801 s, because the model mainly matches the SMPL model with two-dimensional joint points based on parameter optimization, and the optimization process takes a lot of time; SMPL needs to consider not only the basic
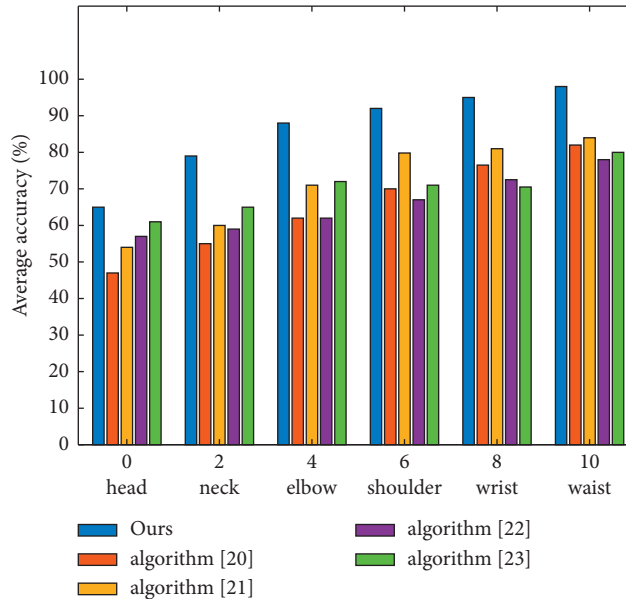
Figure 8: PDJ curve comparing this algorithm with other algorithms on VIPS-VideoPose database.

Table 3: Comparison of PCP value results between this algorithm and other algorithms on vipsvediopose.

Methods	Algorithm (%) [20]	Algorithm (%) [21]	Algorithm (%) [22]	Algorithm (%) [23]	Ours (%)
Neck	78.1	62.8	78.2	57.8	87.1
Elbow	80.9	69.3	79.4	54.7	88.2
Shoulder	91.7	83.3	90.5	34.2	95.1
Wrist	65.4	84.2	93.2	78.5	92.7
Waist	63.8	83.0	92.5	81.0	95.9

human body configuration, but also whether the joint density can control the movement of points. Therefore, when there are no obvious joints in the chest, several joints are added, while the number of joints on the hand is less. For better manikin expression, SMPL adds more points to finer parts, such as fingers. According to the area of normal people, there should be fewer points, but this is unfavorable to SMPL model. The average time per frame of algorithm [21] is 6.281 s. The model combines a large number of 3D attitude prior constraints and 2D data sets, so its attitude estimation has high efficiency; The average running time of each frame of the algorithm in this paper is 1.115 s, in which the process of two-dimensional human posture estimation takes an average of 0.015 s, and the process of three-dimensional SMPL model parameter estimation and three-dimensional joint point calculation takes an average of 1.10 s. The fully connected network structure design proposed in this paper saves a lot of model calculation time, improves the accuracy and maintains high efficiency.

4.3. Results of Comparative Experiments. In order to further illustrate the effectiveness of the algorithm in the field of human posture tracking, the performance of the algorithm is compared with four methods in videopose2.0 and VIPs Videopose databases. The four mainstream methods selected include: (1) the improved method of

graph model in document [21], (2) the model proposed in document [22], (3) the extensible model proposed by document [23], and (4) the hybrid component model proposed in document [24]. In order to ensure the comparability of the experimental results, the training or test of the above four methods are exactly the same as that of the algorithm.

Generally speaking, the PDJ and PCP values of each algorithm when detecting the head are basically close to the real values, so the positioning results of these two joint points are not recorded in this experiment. This experiment only counts the positioning accuracy of 6 joint points and 4 parts (upper and lower arms) on videopose 2.0 and VIPs videopose databases. Figure 8 shows the PDJ curve of this algorithm and the above four methods on VIPs Videopose database. As can be seen from Figure 8, the algorithm adopted in this paper is higher than the conventional algorithm in the positioning accuracy of each joint point, and can obtain better human posture tracking effect. In particular, for the wrist joint points most concerned in the process of human posture estimation, the positioning accuracy of this algorithm is significantly higher than that of the other four methods. For example, taking 30 pixels as the threshold, the wrist PDJ value of this algorithm is 98.3%, which far exceeds the PDJ value of reference [20] 82.1%, reference [21] 84.1%, reference [22] 78.2% and reference [23] 80.1%.

Table 3 shows the comparison results of PCP values of this algorithm and the above four methods (with 30 pixels as the threshold) on VIPs Videopose database. The PCP value of the waist of this algorithm is 95.9%, which far exceeds the PCP value of 63.8% in reference [20], 83.0% in reference [21], 92.5% in reference [22] and 81.0% in reference [23]. It is confirmed that the algorithm still maintains good positioning accuracy on VIPs videopose.

5. Conclusion

In this study, the structure of traditional neural network is improved. In order to study the posture estimation of multiple people in sports competition video, a human posture estimation algorithm based on graph neural network is adopted by using Multi level characteristic diagram. After statistics of 6 joint points and 4 parts on VIPs videopose database show that the graph neural network algorithm can complete the positioning function of human parts, so as to achieve accurate human body tracking effect. Comparing the PDJ curves and PCP values of varIoUs algorithms, it can be seen that the algorithm based on graph neural network greatly improves the positioning accuracy of varIoUs parts of the human body compared with other algorithms. Therefore, when applied to athlete posture tracking in sports competition video, this method can effectively estimate the posture of multiple people.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

Acknowledgments

References

[1] R. A. Minhas, A. Javed, A. Irtaza, M. T. Mahmood, and Y. B. Joo, "Shot classification of field sports videos using AlexNet Convolutional Neural Network," *Applied Sciences*, vol. 9, no. 3, p. 483, 2019.

[2] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-Pose: real-time human skeletal posture estimation using mmWave radars and CNNs," *IEEE Sensors Journal*, vol. 20, no. 17, pp. 10032–10044, 2020.

[3] Q. Dang, J. Yin, B. Wang, and W. Zheng, "Deep learning based 2d human pose estimation: a survey," *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 663–676, 2019.

[4] A. Martínez-González, M. Villamizar, O. Canévet, and O. Jean-Marc, "Efficient convolutional neural networks for depth-based multi-person pose estimation," *IEEE Transactions Circuit System Video*, vol. 30, no. 11, pp. 4207–4221, 2019.

[5] D. Groos, H. Ramampiaro, and E. A. Ihlen, "EfficientPose: scalable single-person pose estimation," *Applied Intelligence*, vol. 51, no. 4, pp. 2518–2533, 2021.

[6] W. Kim, J. Sung, D. Saakes, C. Huang, and S. Xiong, "Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose)," *International Journal of Industrial Ergonomics*, vol. 84, Article ID 103164, 2021.

[7] Y. Pang, Y. Yuan, X. Li, and J. Pan, "Efficient HOG human detection," *Signal Processing*, vol. 91, no. 4, pp. 773–781, 2011.

[8] T. Mordan, N. Thome, G. Henaff, and M. Cord, "End-to-end learning of latent deformable part-based representations for object detection," *Internation Journal of Computer Vision*, vol. 127, no. 11, pp. 1659–1679, 2019.

[9] M. Vasileiadis, C.-S. Bouganis, and D. Tzovaras, "Multi-person 3D pose estimation from 3D cloud data using 3D convolutional neural networks," *Computer Vision and Image Understanding*, vol. 185, pp. 12–23, 2019.

[10] A. Newell, K. Yang, and J. Deng, "Hourglass networks for human pose estimation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 483–499, Amsterdam, The Netherlands, October 2016.

[11] R. Gu, G. Wang, Z. Jiang, and J. Hwang, "Multi-person hierarchical 3d pose estimation in natural videos," *IEEE Transactions Circuit System Video*, vol. 30, pp. 4245–4257, 2019.

[12] Z. Liu, J. Zhu, J. Bu, and C. Chen, "A survey of human pose estimation: the body parts parsing based methods," *Journal of Visual Communication and Image Representation*, vol. 32, pp. 10–19, 2015.

[13] H. J. Bae, G. J. Jang, Y. H. Kim, and J. P. Kim, "LSTM (long short-term memory)-based abnormal behavior recognition using AlphaPose," *KIPS Transactions on Software and Data Engineering*, vol. 10, no. 5, pp. 187–194, 2021.

[14] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu, and J. Sun, "Cascaded pyramid network for multi-person pose estimation," in *Proceedings of the IEEE International Conference Computer vision pattern recognit*, pp. 7103–7112, Salt Lake City, UT, USA, June 2018.

[15] D. Sasakawa, N. Honma, T. Nakayama, and S. Iizuka, "Human posture identification using a MIMO array," *Electron*, vol. 7, no. 3, p. 37, 2018.

[16] S. Payandeh, "A novel depth image analysis for sleep posture estimation," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 5, pp. 1999–2014, 2019.

[17] A. Nadeem, A. Jalal, and K. Kim, "Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy Markov model," *Multimedia Tools and Applications*, vol. 80, no. 14, pp. 21465–21498, 2021.

[18] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2d human pose estimation: new benchmark and state of the art analysis," in *Proceedings of the IEEE Internaional Conferenece Computer vision pattern recognition*, pp. 3686–3693, Columbus, OH, USA, June 2014.

[19] T. Y. Lin, M. Maire, S. Belongie et al., "Microsoft coco: common objects in context," *European conference on computer vision*, pp. 740–755, 2014.

[20] R. Liu, "A study of athlete pose estimation techniques in sports game videos combining multiresidual module convolutional neural networks," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 4367875, 10 pages, 2021.

[21] A. Jalal, I. Akhtar, and K. Kim, "Human posture estimation and sustainable events classification via pseudo-2D stick model and K-ary tree hashing," *Sustainability*, vol. 12, no. 23, p. 9814, 2020.

[22] G. Diraco, A. Leone, and P. Siciliano, "Human posture recognition with a time-of-flight 3D sensor for in-home applications," *Expert Systems with Applications*, vol. 40, no. 2, pp. 744–751, 2013.

[23] K Adhikari, H. Bouchachia, and H. Nait-Charif, "Deep learning based fall detection using simplified human posture," *International Journal of Computer, Information, and Systems Science, and Engineering*, vol. 13, no. 5, pp. 251–256, 2019.

[24] G. S. Chleboun, A. B. Busic, K. K. Graham, and H. A. Stuckey, "Fascicle length change of the human tibialis anterior and vastus lateralis during walking," *Journal of Orthopaedic & Sports Physical Therapy*, vol. 37, no. 7, pp. 372–379, 2007.

Research Article

Real-Time Footprint Planning and Model Predictive Control Based Method for Stable Biped Walking

Song Wang [ID],[1] Songhao Piao [ID],[1] Xiaokun Leng,[1] Zhicheng He,[1] Xuelin Bai,[2] and Li Huazhong[3]

[1]*School of Computer Science, Harbin Institute of Technology, Harbin, China*
[2]*Leju Robotics, Shenzhen, China*
[3]*Shenzhen Institute of Information Technology, Shenzhen, China*

Correspondence should be addressed to Song Wang; song.w@hit.edu.cn

In order to walk in a physical environment, the biped will encounter various external disturbances, and walking under persistent conditions is still challenging. This paper tries to improve the push recovery performance based on capture point (CP) and model predictive control. The trajectory of zero moment point (ZMP) and center of mass are solved and predicted in a limited time horizon. Online footprint generator is combined with MPC walking pattern generation, which can keep biped stable in the next few steps, and projection of ZMP is used to calculate the next footprint and reach the target CP in an incremental way. Verification of the proposed stable biped walking method is conducted by simulation and experiments.

1. Introduction

tBiped robot is one kind of mobile robot and belongs to the field of leg robot. In the world of robotics, biped walking is one of the fields with the most interesting research [1]. Robots need to interact and collide intermittently with the environment to complete specific actions and tasks [2]. Compared with other forms of robots, biped robot has more advantages: the biped with a humanoid structure can adapt to various environments, and it can replace humans in many dangerous jobs and offer services for humans, such as artificial limb and medical rehabilitation [3].

The biped robot model is nonlinear dynamic with multiple degrees of freedom and constraints, and it is important to interpret and predict the long-term behavioral trends of the dynamic system for various control tasks. A simplified biped walking linear inverted pendulum model (LIPM) is the baseline of humanoid walking pattern controller [4, 5], and passive walking and under-actuated robots have certain advantages in energy-saving walking. But passive walking has poor terrain adaptability and can only

walk downhill or on flat ground; while under-actuated walking robots require special structural design and rely on complex walking control strategies [6]. Zhang et al. [7] proposed a walking pattern generator for omnidirectional walking on a slope and uneven plane and realized the oblique walk and heading for any direction, but the performance of push recovery under external persistent force is not analyzed. Oliver et al. proposed an improved model named the flexible linear inverted pendulum model (FLIPM) which integrate damped spring and second center of mass to LIP [8], which realized a stable walk on a servo-driven biped robot. However, the performance is only improved in the z direction, which is not included by FLIPM and causes a rotation of the root alongside the x-axis towards the ankle joint lifting up from the plane. Biped robot has dozens of degrees of freedom and no fixed base, which leads to its dynamic system is very complex. Juang et al. proposed the fully connected recurrent neural network optimized by continuous multiobjective ant colony optimization to form CpG to solve the multiobjective gait generation problem of Nao robot. An open-source multiple DOF servo-driven

robot Robotis op [9] is used to test their method performance, but the coupled oscillator model is not able to keep stability under an unpredictable perturbation.

In the process of walking, the biped robot will inevitably be affected by various uncertain factors, such as the uncertainty of mathematical model based on different assumptions [10], the uncertainty of robot system parameters caused by the structural size, material properties, manufacturing and assembly errors, and the random uncertainty of driving torque caused by actuator noise and joint friction, and the external environment disturbance and plane condition uncertainty in the process of robot walking [11]. Pratt et al. [12] introduced the concepts of capture point and capture area by LIPM and flywheel model, and CP is the key metric for push recovery of the biped which shows the target direction to maintaining body stability by keeping center of pressure (CoP) inside the foot support area. But how the capture region adapts with other biped models is not discussed. Majid et al. generate walking patterns for the biped by two stages: computes the best step location and duration and adapts these values using divergent component of motion (DCM) measurement. Englsberger [13] designed CP following and CP step-end controller and has proved the antipush stability of CP control principle, but real-time footprint position adjustment is not included to enhance the robustness of the walking pattern against external perturbations. Since the large number of joints, it is very complex to use CPG model to describe the joint angle for the 3D humanoid robot. Q-learning-based CPG model [14, 15] can keep the body recover from horizon perturbation, but it need different expertise to redesign the controller. In addition, the motion obtained by this gait control method is generally not optimal, and the anti-interference in vertical direction ability is poor.

Humanoid walking pattern generation can be solved by model predictive control (MPC) with efficient constraint handling [16]. The optimization-based method is utilized so as to generate, at each iteration, the optimal gait trajectory of the system satisfying the given constraints. However, it is well known that the action of a bounded persistent disturbance can destabilize a predictive controller which has been designed to be stabilizing for the nominal case [17]. These controllers can be used with a cost function that at the specific problem of generating trajectory by quadratic programming in linear system control, a fast optimization algorithm is proposed. However, several papers like [18–20] used restrictions on model state at the end of the prediction horizon. Since the state of the system is completely dependent on the beginning and the end of the horizon, motions lack the flexibility to perform secondary control objectives. Amos et al. [21] proposed the strategy architecture designed for end-to-end training, the robot learned to combine high-level planning strategy with low-level motion controller to realize autonomous navigation on a curved path, which is less computationally and memory intensive compared to traditional optimization solutions, but the model used in that article is still linear which is hard to extend to complex systems like the biped walking. Scianca et al. [22] introduced an IS-MPC framework for gait generation which extend LIPM with ZMP trajectory input, and recursive feasibility of internal stability is realized for biped dynamics, but the cost of control and response speed of method are not analyzed for position control based humanoid robot [23].

The gait trajectory planning method based on LIP does not consider the external interference force, so it cannot walk in a contact-rich environment. The gait pattern generator can only calculate the trajectory position at the next time according to the state transition equation but cannot predict the trajectory in the future. This paper proposes footprint planning and MPC method for stable biped walking considering the capture point principle. The framework of the method is summarized in Figure 1. LIPM is used to linearize the dynamic model of biped walking, and MPC is used as a gait pattern generator to generate a smooth trajectory of CoM. Online footstep generator based on capture point feedback is combined with the MPC controller, and the biped can take one or more steps to reach the stable state and realize omnidirectional walking on the horizon plane. And a projection function maps the target ZMP into support polygon which makes inverse kinematics is solvable.

2. Materials and Methods

2.1. Approximation of Walking Model. Approximation acts effective method to deal with complex systems. Most humanoid robots use the model-based walking planning method to get the gait pattern and abstract the biped dynamics equation from the model by approximating the center of mass (CoM). The classical inverted pendulum model is widely used for humanoid walking pattern generation. Some conditions are mentioned first as follows:

(1) Whole-body mass of the robot is concentrated on the CoM.

(2) CoM height h_{CoM} in 3-dimensional space is invariant.

(3) Biped legs are massless and contact with the ground through a pivot that can be rotated.

Under the joint action of the supporting moment and the force along the stretching direction of the rod, the dynamic equation of LIPM is established, and the trajectory of the CoM is obtained by solving the dynamic equation. Using the LIPM, let p denote the position of the center of pressure (CoP) on the plane, and the horizontal dynamics differential equation of CoM is

$$\ddot{x} = \omega_0^2 (x - p), \tag{1}$$

where $w_0 = \sqrt{g/h_{CoM}}$ and g is gravity force, and h_{CoM} is the constant height of CoM along z-axis. x denotes the position of CoM along x-axis, and $\ddot{x}$ denotes body acceleration. This approximation (1) decouples the sagittal and coronal motions of the biped robot, so we will focus on the x-axis motion throughout this paper, and y direction motion is totally identical.
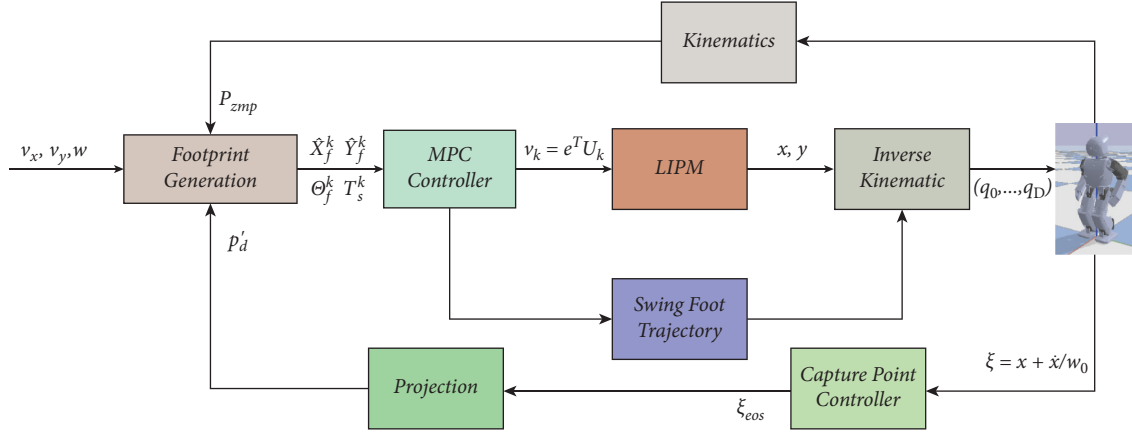
FIGURE 1: Block scheme of the proposed CP- and MPC-based framework for walking gait.

Considering the discretization at a minimum interval time T, CoM, and CoP are discretized. Let input control signal $u(t) = \dddot{x}$ which is the jerk of CoM position in state-space equation, then the system state transition at time $t = kT, k = 1, 2, \ldots$ with notation

$$\widehat{x}_k = \begin{bmatrix} x(kT) \\ \dot{x}(kT) \\ \ddot{x}(kT) \end{bmatrix}, u_k = \dddot{x}(kT), p_k = p(kT). \tag{2}$$

From (2), we can get the ZMP position

$$p_k = \begin{bmatrix} 1 & 0 & 1/w_0^2 \end{bmatrix} \widehat{x}_k. \tag{3}$$

For the system state $\widehat{x}_k$, the discretized system of LIPM is

$$\begin{aligned} \widehat{x}_{k+1} &= \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \widehat{x}_k + \begin{bmatrix} T^3/6 \\ T^2/2 \\ T \end{bmatrix} u_k \\ &= A\widehat{x}_k + Bu_k. \end{aligned} \tag{4}$$

In the process of bipedal walking, the position of ZMP p_k always falls into the support polygon of the foot, so the range of ZMP can be limited within

$$p_k^{\min} < p_k < p_k^{\max}. \tag{5}$$

The range of ZMP $[p_k^{\min}, p_k^{\max}]$ depends on foot contact condition with the horizontal plane at time kT. According to the state transition (2), the next state at time $k + 1$ is decided by the current state x_k and input control u_k. The core idea of MPC makes the actual ZMP p_k most likely close to reference ZMP trajectory p_k^{ref} at the minimum control cost at time kT. Then, the optimal control sequence $u_k, u_{k+1,\ldots}$ can be get by solving the quadratic problem (QP).

$$\min_{u_k, u_{k+1}, \ldots} \sum_{i=k}^{\infty} \frac{1}{2} \alpha \left(p_{i+1} - p_{i+1}^{ref} \right)^2 + \frac{1}{2} \beta u_i^2, \tag{6}$$

where $\alpha/\beta \, (\alpha > 0, \beta > 0)$ are factors to balance the system response ratio and control cost. If α/β rises, the tacking speed of system and control costs increase. And if α/β is lower, the tracking speed of the system and control costs decrease.

2.2. Capture Point Controller.
For push resistance of the bipedal, instantaneous capture point (ICP) was utilized. During the motion of a LIP model, there exists a metric point on a walking plane, where placing the center of the landing foot can stabilize the motion of the model. This point can be calculated as

$$\xi = x + \frac{\dot{x}}{w_0}, \tag{7}$$

where ξ is the x or y component of ICP position on plane. Due to the limitation of robot dynamics, when the capture point is not captured, the CP will be away from CoP in a straight line until it falls down that is noted by Figure 2.

The features of the capture point can be described as follows:

(1) Due to the regular dynamics of the robot, if the point is not captured, with time, the ICP moves in a straight line joining the CoP and the CP in a direction away from the foot, as shown in Figure 2.

(2) Zero-step capturability: If the ICP falls inside the support polygon, then the robot can balance itself with the application of a good control like MPC; however, if the ICP falls outside the support polygon, the robot will have to take at least one step to protect itself from falling down.

(3) One-step capturability: The model can come to stability within a step if the ICP falls within the reachable range. The reachable region is bounded by the maximum step length.

According to equation (1) and equation (7), the dynamic relationship between CP and ZMP is

$$\dot{\xi}_x = w_0 \left(\xi_x - p_x \right). \tag{8}$$

Since the pole of transition function is w_0, equation (7) is an unstable system without external additional input. The idea of CP control is to generate the ZMP trajectory so that the current CP reaches the target CP within a given time interval, and then, the CoM follows the planning CP curve. The solution of equation (7) in the time domain is
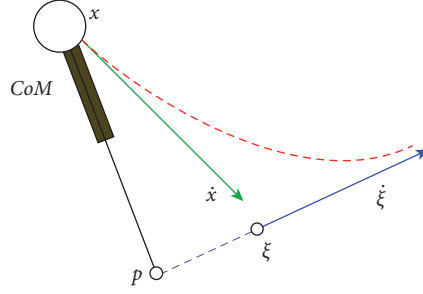
FIGURE 2: Top view of the LIPM with point foot for a given initial state.

$$\xi_x(t) = e^{\omega_0 t}\left(\xi_{x,0} - p_x\right) + p_x, \tag{9}$$

where $\xi_{x,0}$ is the initial position of CP. It can be seen that when ZMP p_x is constant, CP will change exponentially. Let $\xi_{x,d}, \xi_x$ represent the target CP and current CP, respectively, and time from $\xi_{x,d}$ to ξ_x is dT, and then, (9) can be discretized as

$$p_x = \frac{\xi_{x,d} - e^{w_0 dT}\xi_x}{1 - e^{w_0 dT}} = \frac{1}{1 - e^{w_0 dT}}\xi_{x,d} - \frac{e^{w_0 dT}}{1 - e^{w_0 dT}}\xi_x. \tag{10}$$

The goal of end-of-step CP controller: ZMP trajectory is generated by $\xi_{x,d}$ and variable time span dT. Let $\xi_{x,eos}$ represent the end of step at the ending of each foothold, and $\xi_{x,d} = \xi_{x,eos}$. Suppose the center of the support foot is approximately regarded as the ZMP point p_i at step i, and the initial CP is

$$\xi_{init,i} = p_i + \left(\xi_{eos,i} - p_i\right)/e^{w_0 t_{step}}, \tag{11}$$

where t_{step} is the duration of each step. Iterative relation of CP state is $\xi_{eos,i-1} = \xi_{init,i}$. When the robot is subjected to external thrust, CP changes suddenly, and the biped needs to replan the landing footprint so that the CoM is adjusted to ξ_{eos}. According to the iterative relationship (11), the CP position in each support leg cycle can be calculated. Then, the system reaches a stable state following the new reference ZMP sequence generated by the CP controller.

ZMP reference trajectory is a continuous curve and constrained by robot dynamics of support polygon. When the system is disturbed by a large external force, CP varies fast out of feet support polygon on the plane. If p_x is calculated by (11), $\hat{p}_d$ might locate out of support polygon. And the incremental progressive method can be adopted for ZMP, and CoM is controlled to move along the target CP direction within the current support polygon. Formula (9) is differentiated as follows:

$$d\xi_{eos} = \frac{\partial \xi_{eos}}{\partial p_d} = \left(1 - e^{w_0 dT}\right)dp_d,$$

$$\Delta\xi_{eos} = \int d\xi_{eos}dt. \tag{12}$$

A method to implement the incremental integration (12) is the ZMP projection as shown in Figure 3. If the projection is not applied, p_d from (11) could beyond the support
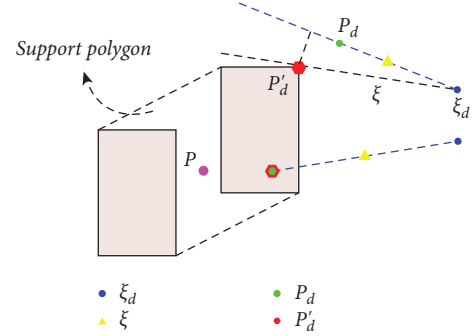
FIGURE 3: Projection of p_d in support polygon.

polygon, and the CoM trajectory generated by p_d will cause falling on the ground.

As Figure 3 shows, we found a new projection ZMP p_d' in support polygon to make CoM follow the target direction of CP. The projection rules can be summarized as follows:

   (i) If p_d is located in support polygon, the target ZMP p_d is in the reachable region of the biped, then projection ZMP $p_d' = p_d$.

   (ii) If p_d is located outside the support polygon, p_d' is the closest point on the inner boundary of the support polygon to $(\xi_d - \xi)$.

In practical application, the number of reverse iterative footprint calculation F is limited. Based on the current foothold p_i and initial CP $\xi_{init,i}$, the position of ZMP can be solved on a limited horizon F. CoM mapping is used to make the CoM approach the target position. But as 3 shows, there is a discontinuity between current p and $p_d' = p_d$ which leads to discontinuous input to the walking pattern generator. To compensate for this problem, the MPC controller is introduced in the next section to track the CP trajectory on the one hand and export smooth CoM trajectory, so that the biped will not lose stability due to the sudden change of dynamic parameters in the process of walking.

2.3. Footprint Generator with Variable Velocity.
The footprint generator module and MPC controller module are implemented synchronously. The input of the footprint generator is reference velocity v_x, v_y and angular velocity w around the z-axis. The output of this generator is the next

footprint position, direction, and time. With the change of input reference velocity, the cycle and step length of biped walking will be affected, and the footprint generator calculates the next foothold position by solving the quadratic optimization problem. The time and orientation of the foothold of the output of the footprint generator will be used as the reference input of the MPC controller in the next step.

As shown in Figure 1, the input of the footprint generator is v_x, v_y, w from time t_k to $t_{k+P} = t_k + T_p$, where T_p is the duration of the footstep. The output of the footprint generator is the sequence $(\widehat{X}_f^k, \widehat{Y}_f^k, \Theta_f^k)$ on time sequence $\mathcal{T}_s^k$, and

$$
\begin{aligned}
\widehat{\mathbf{X}}_f^k &= \left(x_f^1 ... x_f^F \right)^T, \\
\widehat{\mathbf{Y}}_f^k &= \left(y_f^1 ... y_f^F \right)^T, \\
\Theta_f^k &= \left(\theta_f^1 ... \theta_f^F \right)^T, \\
\mathcal{T}_s^k &= \left\{ T_s^1, ..., T_s^F \right\},
\end{aligned}
\tag{13}
$$

where F is the number of next footprints, and $(x_f^j, y_f^j, \theta_f^j)$ is the position and orientation of j-th footprint. T_s^j is the duration from $j-1$-th to j-th footprint. Since the input velocity is variable, the prediction horizon of time T_s is constant, and $\sum_{j=1}^{F} T_s^j = T_p$, and then, distribution of $(x_f^j, y_f^j, \theta_f^j)$ is nonequidistant during T_p.

Assume that the biped is within the j-th support phase switching duration T_s^j, $\overline{v}$ is average velocity, $\overline{T}_s$ is step period duration, and $\overline{L}_s$ is stride length, so $\overline{v} = \overline{L}_s / \overline{T}_s$. Walking velocity is limited by the biped dynamics such as CoM height, degree of freedom, and leg length. $\overline{v}$ is determined by both T_s^j and $\overline{L}_s$. Δv is velocity small variation in Δt, and then,

$$
v = \overline{v} + \Delta v = \frac{\overline{L}_s + \Delta L_s}{\overline{T}_s - \Delta T_s},
\tag{14}
$$

where $\Delta L_s = \gamma \Delta T_s$, and (14) can be expressed as

$$
T_s = \overline{T}_s \frac{\gamma + \overline{v}}{\gamma + v}.
\tag{15}
$$

Let $\overline{v} = 0.01 m/s$, $\overline{T}_s = 0.8 s$, and the relationship between velocity and duration can be shown in Figure 4.

During a small time interval δt, velocity can be approximately regarded as a linear variation. Therefore, the orientation angle of the robot can be approximately constant, and angular velocity is ignored in (14).

Considering time sequence $t_1^j, t_2^2, ..., t_s^F$ iterative relation, the time iterative equation within $[t_k, t_k + T_p]$ is

$$
t_s^j = t_s^{j-1} + \overline{T}_s \frac{\gamma + \overline{v}}{\gamma + v \left(t_s^{j-1} \right)},
\tag{16}
$$

where t_s^0 is the end time of the last support leg duration, and the iteration is over when $j > k + P$. Then, a time sequence $\mathcal{T}_s^k = \left\{ T_s^1, ..., T_s^F \right\}$ is outputted with $T_s^j = t_s^{j+1} - t_s^j$.

According to the omnidirectional motion model of a mobile robot, the bipedal moves to the target point at any angle and direction in the horizontal plane with v_x, v_y, w.
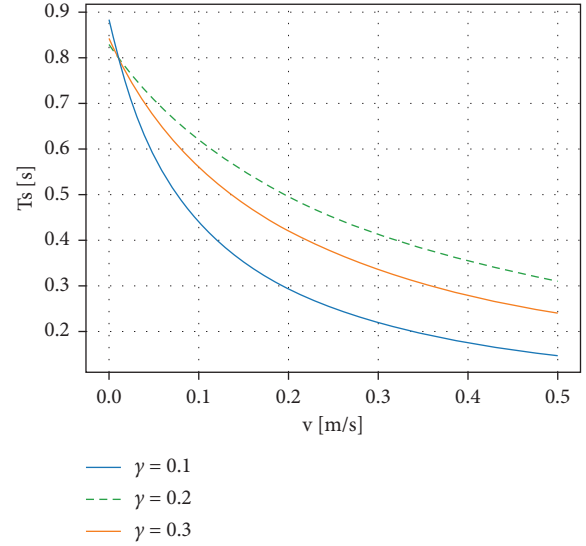
FIGURE 4: Rules for step duration T_s and velocity $\overline{v}$, and comparison of different γ.

$$
\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ \omega \end{pmatrix}.
\tag{17}
$$

The idea of the footprint generator is to allocate the future foothold under the condition of considering the constraints of kinematics and dynamics parameters of the biped in time $\mathcal{T}_s^k = \{T_s^1, ..., T_s^F\}$. When the body root is subjected to push forces, according to (7), $\dot{x}$ will change suddenly, which will affect the planned trajectory of x. The velocity and acceleration change suddenly, which affect the planning trajectory of CoM. The position of the foothold need be adjusted to make the CoM as close as possible near new CP so that the CoM can keep a gradually stable state.

The footprint planning can be transformed into two QP optimization problems as follows:

$$
\left\{ \min_{\Theta_f^k} \sum_{j=1}^{F} \left(\theta_f^j - \theta_f^{j-1} - \int_{t_s^{j-1}}^{t_s^j} \omega(\tau) d\tau \right)^2 st. \left| \theta_f^j - \theta_f^{j-1} \right| \leq \theta_{\max},
\right.
\tag{18}
$$

where $\theta_{\max}$ is the maximum steering angle allowed for two consecutive footprints according to the physics dynamics of the biped. And the second QP problem is

$$
\min_{\widetilde{x}_f^k, \widetilde{r}_f^k} \sum_{j=1}^{F} \left(\widehat{x}_f^j - \widehat{x}_f^{j-1} - \Delta x^j \right)^2 + \left(\widehat{y}_f^j - \widehat{y}_f^{j-1} - \Delta y^j \right)^2.
\tag{19}
$$

where the supporting foot position is $(\widehat{x}_f^0, \widehat{y}_f^0)$ at the starting time t_k. The output of the footprint generator $(X_f^k, Y_f^k, \Theta_f^k)$ is the input of the MPC controller, and the final ZMP and CoM trajectory is calculated by MPC. Figure 5 shows an example of footprint generation where the orientation of the robot coincides with the tangent of the motion path. The swing foot trajectory planning is simple, and in order to avoid sudden changes in the velocity of end point, we use Bezier curves to generate swing foot trajectory and then solve the joints rotation by inverse kinematics.
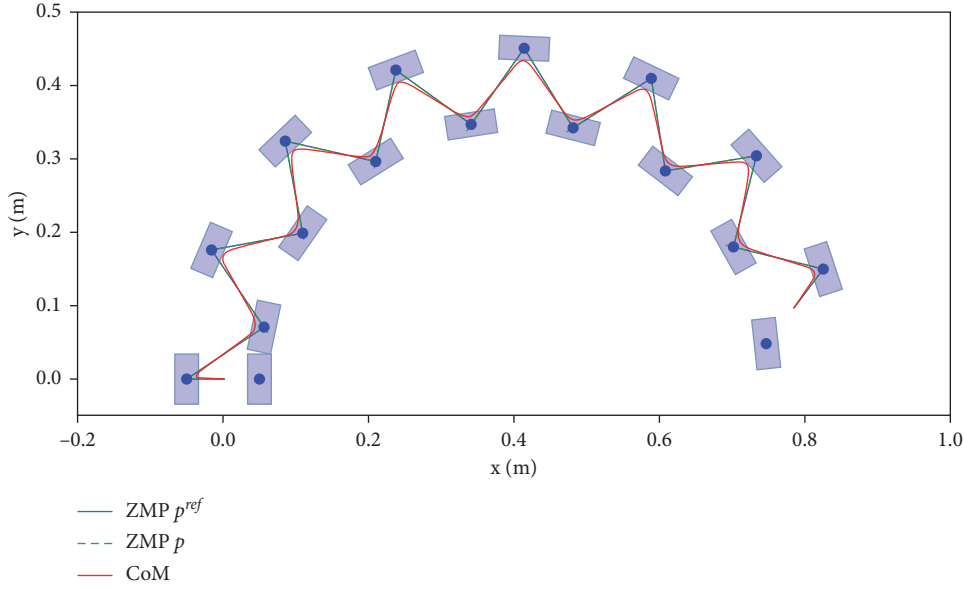
FIGURE 5: Candidate footsteps generated by the footprint generator in half circle demonstrate the omnidirectional feature.

2.4. MPC Controller. Predictive control approximates the robot to a three-dimensional linear inverted pendulum, introduces the state space method, optimizes the performance index by the linear quadratic regulator, and limits the movement of the CoM in the horizontal plane with constant height. The state feedback gain and predictive gain of the predictive controller are generated by solving discrete algebraic Riccati equations. Due to the consideration of state feedback, error feedback, and future target information, a stable walking mode can be generated with less computation.

The system only performs prediction of MPC controller at kT, then measures the actual state $x, \dot{x}$ of the system as feedback by LIPM, and calculates the optimal control by solving QP. N represents the prediction horizon of the MPC controller, and equation (6) can be rewritten within $[kT, (k + N)T]$.

$$\min_{u_k,\ldots,l_{k+N}} \sum_{i=k}^{k+N-1} \frac{1}{2}\alpha\left(p_{i+1} - p_{i+1}^{ref}\right)^2 + \frac{1}{2}\beta u_i^2. \tag{20}$$

The above equation (20) is solved by solving complex algebraic Riccati equation in [5], but this method has low computational efficiency. Within MPC prediction horizon, equation (4) can be transformed to matrix operation. The state transition equation is rewritten with length N as

$$\begin{bmatrix} p_{k+1} \\ \vdots \\ p_{k+N} \end{bmatrix} = \begin{bmatrix} 1 & T & T^2/2 - 1/w_0^2 \\ \vdots & \vdots & \vdots \\ 1 & NT & N^2T^2/2 - 1/w_0^2 \end{bmatrix} \hat{x}_k$$

$$+ \begin{bmatrix} T^3/6 - T/w_0^2 & 0 & 0 \\ \vdots & \ddots & 0 \\ \left(1 + 3N + 3N^2\right)T^3/6 - T/w_0^2 & \cdots & T^3/6 - T/w_0^2 \end{bmatrix}$$

$$\times \begin{bmatrix} u_k \\ \vdots \\ u_{k+N-1} \end{bmatrix}, \tag{21}$$

And (21) can be abbreviated as

$$P_{k+1} = Q_x \hat{x}_k + R_u U_k. \tag{22}$$

According to (20), the QP equation can be represented as

$$\min_{x_k} \frac{1}{2}\alpha\left(P_{k+1} - P_{k+1}^{ref}\right)^2 + \frac{1}{2}\beta U_k^2, \tag{23}$$

And a standard QP form can be generated by substituting (22) into (23)

$$\min_{U_k}\left[\frac{1}{2}\left(\alpha R_u^2 + \beta\right)U_k^2 + \alpha\left(Q_x\hat{x}_k R_u - P_{k+1}^{ref}R_u\right)U_k\right]. \quad (24)$$

According to the definition of MPC, u_k at the current moment is the first prediction of MPC prediction trajectory; that is,

$$u_k = e^T U_k, \quad (25)$$

where $e^T = [1, 0, ..., 0]^T$, and according to the state formula of LIPM (4), the trajectory of x and output position of ZMP at the next moment are calculated as

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k = A\hat{x}_k + Be^T U_k. \quad (26)$$

Within the predictive horizon NT, a balance between response speed and control cost of system can be found by adjusting $\alpha/\beta\,(\alpha > 0, \beta > 0)$, and in this article, $\alpha/\beta = 10^{-5}$. When the system and environment models are well known, NT can approach infinity. But LIPM model itself is a linear approximation to the walking dynamics of biped robot with systematic model errors, and environment cannot be accurately modeled. So prediction horizon could not be infinite. In fact, when N increases, the gain of the system is close to 0, which means that the influence of the predicted future value on the stability of the current system is reduced, but it will significantly decrease the computational efficiency of QP calculation. Therefore, this paper sets the MPC control horizon to $NT = 1s$.

3. Results

The performance of the proposed dynamic balance control framework in Figure 1 is proved by simulation experiment, we present simulation results using the LIPM and MPC simulation in bullet environment, the input velocity is constant, the single leg support phase time is constant $t_{sup} = 0.5s$, supporting leg switching time $t_{sw} = 0.33s$, MPC horizon $N = 30$, and the minimum control frequency is 100 Hz. The external push force $f_{ext} = 0$, and the robot is not subject to external interference, so that walks along a straight line with $N_{step} = 6$, and according to the foothold generator, the robot's foothold position is shown in Figure 6.

Since the LIPM is decoupled from each other in X and Y directions, body CoM and ZMP are solved separately by the MPC controller, as is shown in Figure 7. Although there is slight noise jitter in the output ZMP trajectory, the CoM maintains smooth transition, and there is no big speed mutation when the support legs are exchanged. Therefore, the MPC method has good tracking performance.

For the omnidirectional walking, Figure 5 shows velocity direction of walking can be changed in real time by the footprint generator, and $\theta_{max} = \pi/4$. The robot walks around the center of the circle along with a radius of 2m. The system receives high-level control input v_x, v_y, w, and the optimization problems (18)and (19) is solved at each step. By
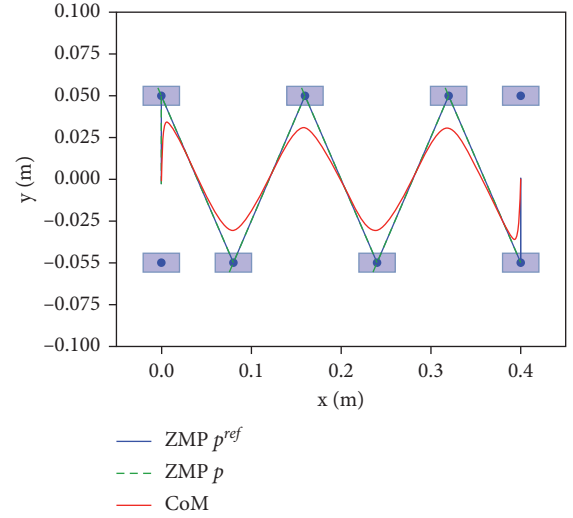
FIGURE 6: The CoM and ZMP trajectory when walking in a straight line $N_{step} = 6$.
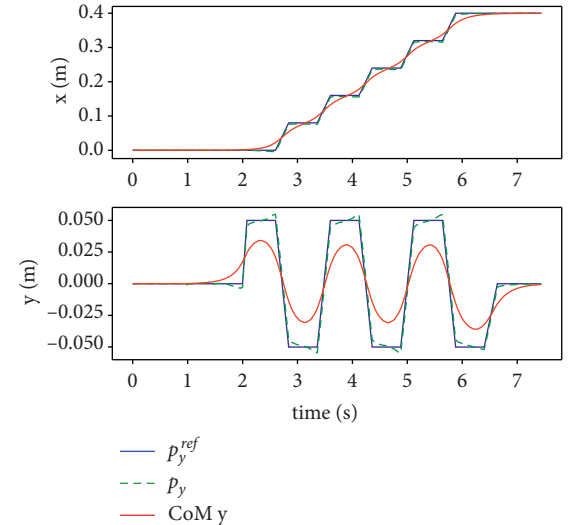
FIGURE 7: When walking in a straight line, the CoM in the X direction (up) and Y direction (down) and ZMP trajectory.

solving the inverse kinematics, the next step foothold rotation is realized by the hip joint rotation; thus, the omnidirectional walking on the horizon plane of the biped is implemented.

The performance of the proposed method is measured in case of external disturbance in x-y plane, as shown in Figures 8–11. An instantaneous external thrust $f_{ext} = 10N$ is added to the CoM of the robot between 4s and 5s, that duration is 0.01 s. It can be seen that the system can adjust the foothold position in 1 step in both directions and be stabilized.

Since the interference of external thrust, the biped will deviate from the predetermined trajectory, but the velocity and direction of biped walking can be changed in real time, and the system has the ability to adjust the input parameters to make the biped walking towards the target position.
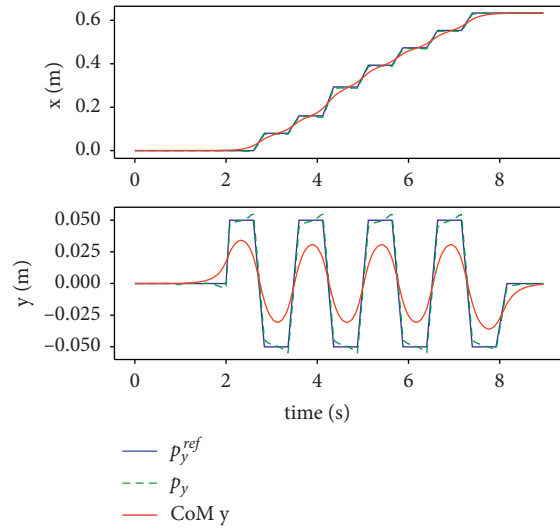
FIGURE 8: Simulation of tracking the trajectory of time with disturbance force in the sagittal direction.
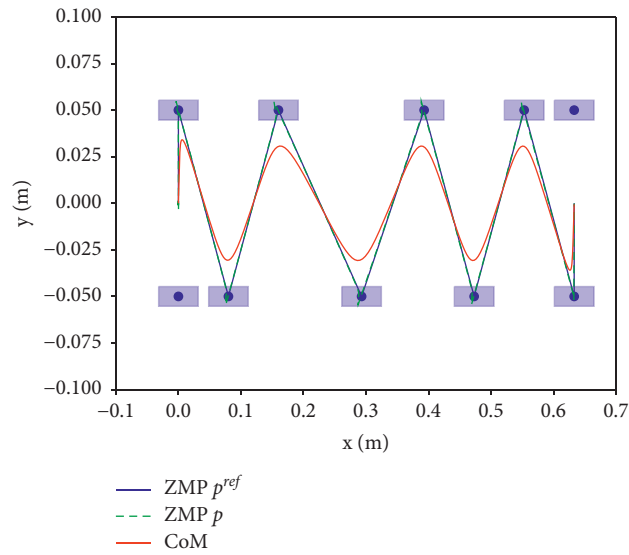
FIGURE 9: Simulation of tracking the trajectory in horizon plane with disturbance force in the sagittal direction.
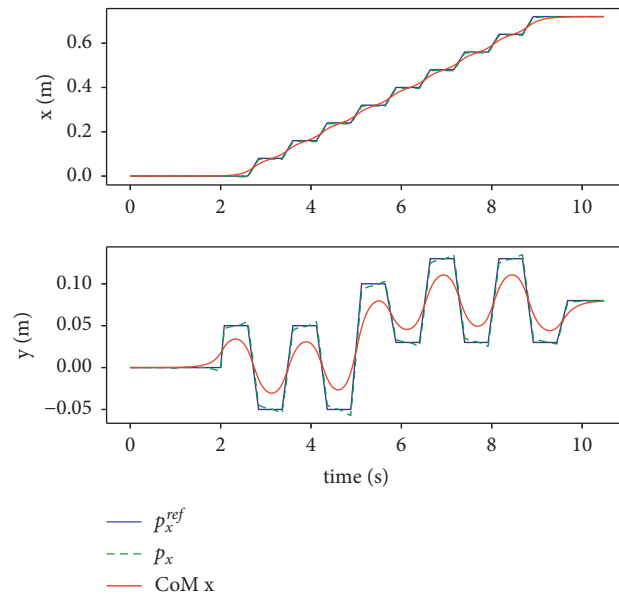
FIGURE 10: Simulation of tracking the trajectory of time with disturbance force in the coronal direction.
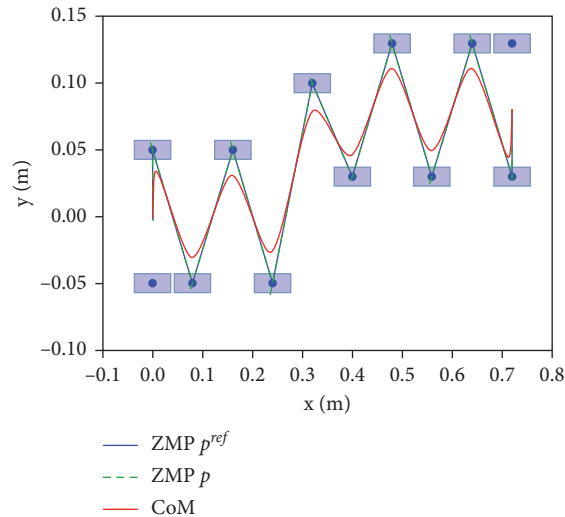
FIGURE 11: Simulation of tracking the trajectory in horizon plane with disturbance force in the coronal direction.

4. Conclusions

At the first, a simplified LIPM is used to linearize the dynamic model of biped walking, MPC is used as a gait pattern generator to keep dynamics and kinematics feasibility, and CoM and ZMP trajectory of the biped is solved and predicted in a limited horizon. Since the bipedal system is an unstable system, it is easy to be disturbed by external thrust, resulting in system instability and falling down. Therefore, this paper designs a bipedal balance control method considering capture point with MPC controller, which realizes the footprints planning with variable speed, so that the robot can walk in omnidirectional directions on the plane. At the same time, the capture point parameters of the system are introduced as feedback. When the system is disturbed by external thrust, the acquisition point will change suddenly. By controlling the ZMP point trajectory, the centroid can track the direction of the capture point, and the system will walk to keep it stable. However, CP may fall outside the support polygon. We use a projection method to make the next landing footprint always fall in the range of the biped reachable area and reach the target CP in an incremental manner. For solving the problem of ZMP input point discontinuity caused by projection, the system uses the MPC method to predict the smooth trajectory curve of centroid in the future according to the new foothold position. The position of joint of the full body is calculated by inverse kinematics and drives the biped realized stable walking.

We only focus on the dynamic walking of bipedal on the horizontal plane. It is necessary to analyze the dynamic walking of bipedal on the nonstationary ground and extend the planning of foothold in a two-dimensional plane to three-dimensional space. On the other hand, the gait pattern generator is mainly a linear inverted pendulum model, which is an approximation of bipedal walking and can not plan complex action behavior. And there will be some improvements in the system approximation model, and a more accurate dynamic and environmental model will be used to predict and analyze the robot's future actions and behaviors.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] C.-F. Juang and Y.-T. Yeh, "Multiobjective evolution of biped robot gaits using advanced continuous ant-colony optimized recurrent neural networks," *IEEE Transactions on Cybernetics*, vol. 48, no. 6, pp. 1910–1922, 2017.

[2] P.-B. Wieber, R. Tedrake, and S. Kuindersma, "Modeling and control of legged robots," *Springer Handbook of Robotics*, pp. 1203–1234, 2016.

[3] X. Yang, H. She, H. Lu, T. Fukuda, and Y. Shen, "State of the art: bipedal robots for lower limb rehabilitation," *Applied Sciences*, vol. 7, no. 11, p. 1182, 2017.

[4] S. Kajita, K. Fumio, K. Kenji, Y. Kazuhito, and H. Hirohisa, "The 3D linear inverted pendulum mode: a simple modeling for a biped walking pattern generation,"vol. 1, pp. 239–246, in *Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, pp. 239–246, IEEE, Maui, HI, USA, 29 Oct.-3 Nov. 2001.

[5] J. Park and Y. Youngil, "General ZMP preview control for bipedal walking," in *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, pp. 2682–2687, IEEE, Rome, Italy, 10-14 April 2007.

[6] J.-w. Luo, Y.-l. Fu, and S.-g. Wang, "3D stable biped walking control and implementation on real robot," *Advanced Robotics*, vol. 31, no. 12, pp. 634–649, 2017.

[7] Z. Yu, X. Chen, Q. Huang et al., "Gait planning of omnidirectional walk on inclined ground for biped robots," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 7, pp. 888–897, 2016.

[8] L. Lanari, U. Oliver, H. Seth, and S. Ingmar, "Boundedness approach to gait planning for the flexible linear inverted pendulum model," *In Robot World Cup*, pp. 58–70, 2016.

[9] X. Li, L. Yangmin, and C. Xinzhe, "Kinematic analysis and gait planning for a Darwin-OP Humanoid Robot," in *Proceedings of the In 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1442–1447, IEEE, Qingdao, China, 3-7 Dec. 2016.

[10] H.-M. Joe and J.-H. Oh, "Balance recovery through model predictive control based on capture point dynamics for biped walking robot," *Robotics and Autonomous Systems*, vol. 105, pp. 1–10, 2018.

[11] H. Gritli and S. Belghith, "Identification, stability and stabilization of limit cycles in a compass-gait biped model via a hybrid poincaré map," *Advances and Applications in Nonlinear Control Systems*, pp. 259–289, 2016.

[12] J. Pratt, C. John, D. Sergey, and G. Ambarish, "Capture point: a step toward humanoid push recovery," in *Proceedings of the In 2006 6th IEEE-RAS international conference on humanoid robots*, pp. 200–207, IEEE, Genova, Italy, 4-6 Dec. 2006.

[13] J. Englsberger, O. Christian, A. R. Máximo, A.-S. Alin, and H. Gerhard, "Bipedal walking control based on capture point dynamics," in *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4420–4427, IEEE, San Francisco, CA, USA, 25-30 Sept. 2011.

[14] D. Tran, H. Fred, and N. John, "A humanoid robot learns to recover perturbation during swinging motion," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3701–3712, 2018.

[15] M. Missura and B. Sven, "Gradient-driven online learning of bipedal push recovery," in *Proceedings of the In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 387–392, IEEE, Hamburg, Germany, 28 Sept.-2 Oct. 2015.

[16] M. Parsa and F. Mohammad, "Robust nonlinear model predictive trajectory free control of biped robots based on nonlinear disturbance observer," in *Proceedings of the 2010 18th Iranian Conference on Electrical Engineering*, pp. 617–622, IEEE, Isfahan, Iran, 11-13 May 2010.

[17] F. M. Smaldone, S. Nicola, M. Valerio, L. Leonardo, and O. Giuseppe, "Gait generation using intrinsically stable MPC in the presence of persistent disturbances," in *Proceedings of the 2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)*, pp. 651–656, IEEE, Toronto, ON, Canada, 15-17 Oct. 2019.

[18] H. Diedam, D. Dimitar, W. Pierre-Brice, M. Katja, and D. Moritz, "Online walking gait generation with adaptive foot positioning through linear model predictive control," in *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1121–1126, IEEE, Nice, France, 22-26 Sept. 2008.

[19] A. Herdt, P. Nicolas, and W. Pierre-Brice, "Walking without thinking about it," in *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 190–195, IEEE, Taipei, Taiwan, 18-22 Oct. 2010.

[20] A. Herdt, D. Holger, W. Pierre-Brice, D. Dimitar, M. Katja, and D. Moritz, "Online walking motion generation with automatic footstep placement," *Advanced Robotics*, vol. 24, no. 5–6, pp. 719–737, 2010.

[21] B. Amos, D. J. R. Ivan, S. Jacob, B. Byron, and J. Zico Kolter, "Differentiable mpc for end-to-end planning and control," arXiv preprint: http://arXiv.org/abs/1810.13400, 2018.

[22] N. Scianca, D. De. Simone, L. Lanari, and G. Oriolo, "MPC for humanoid gait generation: stability and feasibility," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1171–1188, 2020.

[23] C.-F. Juang and Y.-T. Yeh, "Multiobjective evolution of biped robot gaits using advanced continuous ant-colony optimized recurrent neural networks," *IEEE Transactions on Cybernetics*, vol. 48, no. 6, pp. 1910–1922, 2018.

Research Article

Spatiotemporal Diurnal Modulation Characteristic of Wind Speed and Power Generation Revealed by Its Measured Data Processing

Jie Wan [iD],[1,2] **Kun Yao** [iD],[1] **Guorui Ren,**[3] **Ke Han,**[4] **Qi Wang,**[2] **and Jilai Yu**[2]

[1]*Laboratory for Space Environment and Physical Sciences, Harbin Institute of Technology, Harbin 150001, China*
[2]*School of Electrical Engineering &Automation, Harbin Institute of Technology, Harbin 150001, China*
[3]*The State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China*
[4]*School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China*

Correspondence should be addressed to Kun Yao; 19b353009@stu.hit.edu.cn

Received 12 January 2022; Accepted 19 March 2022; Published 31 March 2022

Academic Editor: Akshi Kumar

Atmospheric turbulence is an intrinsic factor that causes uncertainty of wind speed and its power generation by wind turbine. The research of atmospheric turbulence characteristics of wind farms can be used to reduce this uncertainty. In this paper, enough measurement data getting from actual wind farms is used for information processing to quantitatively analyze the daily variation of wind speed and its power output characteristics. Furthermore, the concept of spatiotemporal diurnal modulation characteristics of atmospheric turbulence is proposed with a global scope, which is an intrinsic property of wind. Besides the daily variation characteristics, the average hourly wind speed has a short-term modulation effect on its turbulence and provides a modulation characteristic on wind speed uncertainty. Moreover, the long-term modulation process is affected by seasonal and regional factors, indicating that it has spatiotemporal characteristics. This atmospheric turbulence characteristic has similar effects on characteristic description parameters. However, the characteristics description parameters of wind speed and wind power variation fail to reflect such intrinsic characteristics that are not affected by the spatiotemporal diurnal modulation characteristics of atmospheric turbulence. This indicates that they do not have diurnal characteristics. Finally, a time-varying model combined with the spatiotemporal diurnal modulation characteristics of wind speed and its power generation is discussed by applying on the evaluation of frequency control in power systems. It is shown that the results obtained by measured data processing could improve the power generation quality of large-scale wind power effectively.

1. Introduction

As one of the renewable resources, wind energy is being used worldwide. Researchers from various countries have carried out lots of work on the current status and future of wind power in their countries [1–6]. Development of wind power from the viewpoints of political, social, and technical issues was analyzed in [7]. In China, wind power is the leading energy development sector under the low-carbon development policy, while a series of wind power development plans were drawn up by the government [8–10].

However, the uncertainty in wind speed poses challenges for the utilization of wind power [11–13]. In addition to randomness and volatility, intermittency is another problem that plagues the large-scale application of wind energy [14]. A series of strategies and paths covering the source sides, grid sides, and load sides have been studied and explored to use large-scale wind power safely and efficiently [2, 5, 15]. In [16], China's feed-in tariff mechanism for large-scale wind power is shown. Moreover, the characteristics of Chinese energy structure determine the developmental path of rapid and deep peak regulation of thermal power [17, 18]. Several novel methods were proposed for wind power to smooth the output power of wind energy [19, 20].

Considering the uncertainty of wind energy on grid connection, research work focuses on wind power

prediction and how to do global dispatch planning and design by prediction results [6, 21, 22], including data model, physical model, and hybrid model [23–25]. The continuous development of artificial intelligence algorithms has further expanded the scope of its application to new areas [26, 27]. Using data obtained from data-rich farm, high-dimensional data features are obtained through a series of data extraction methods and applied to newly built farm. This transfer learning research is attracting further attention [28].

However, prediction error cannot be avoided due to wind speed uncertainty [29]. Further, quantitative characterization of the wind speed and generated power's persistence and variation is also urgently needed [30]. In view of the physical nature of the random fluctuation of wind speed and the characteristic of wind power, several studies have been carried out on their uncertainty, including instantaneous characteristics (power spectrum), short-period characteristics (daily variation), and long-period statistical characteristics.

In terms of the nature relationships of range to standard deviation of wind fluctuations, the fluctuation variance caused by turbulence is dependent on the mean wind speed per hour [31]. Literature [32] tries to predict the turbulence standard deviation of wind speed. Because of the importance of long-period characteristics, the spatiotemporal complementarity between solar and wind power in the Iberian Peninsula has been researched as the key problem [33, 34]. Additionally, wind speed variance is also an important parameter to characterize wind and turbulence intensity. Turbulence intensity can be used in fan safety design [35–37], life analysis [38–41], and wind farm layout design [42, 43] and, hence, has long been the focus of many researchers.

Although characterization of the wind resource is important [44], the most significant one is the intermittency caused by the inherent instability of atmospheric turbulence. In [45, 46], the measurement and analysis of intermittency for power generation and wind speed were performed based on historical data of wind speed and wind power. Although the intermittent characterization index of wind speed and power has the characteristics of daily cycle, neither the practical application of daily cycle nor the global perspective on the summarized daily modulation has been discussed.

In fact, due to the existence of the diurnal modulation of atmospheric turbulence, the diurnal periodic characteristics are reflected in the average wind speed and temperature [47, 48]. Many factors, such as atmospheric stability [49], time scale [48], near-surface temperature, and turbulence intensity [50], will all influence the wind characteristics. Therefore, modeling and research on diurnal cycle characteristics are also being carried out [51, 52].

In summary, the average wind speed followed by a certain law has been deeply studied presently. Scholars have studied several systematic algorithms on wind speed and its generated power prediction. However, only few studies study on turbulent wind speed due to the uncertainty caused by turbulence, and only a few researchers have conducted studies on turbulence intensity from the application aspects of wind turbines' life analysis and design of wind farm

layout. Although several researches have pointed that the average wind speed, turbulence intensity, and temperature show diurnal characteristics, there is no study that has proposed the intrinsic characteristics of spatiotemporal diurnal modulation of atmospheric turbulence from a global perspective. Furthermore, there is also absence of a systematic analysis of the influences of diurnal modulation on wind speed and wind turbine randomness, volatility, and intermittency based on the essential characterization parameters of its output power.

In this paper, the spatiotemporal diurnal modulation characteristics of atmospheric turbulence and their influence on the indeterminacy of wind speed and power generation are discussed. The rest of this paper is organized as follows: in Section 2, the spatiotemporal diurnal modulation characteristics of atmospheric turbulence are proposed from a global perspective based on the analysis of its physical mechanism. In Section 3, the influence of spatiotemporal diurnal modulation of atmospheric turbulence on wind speed and its power production's random fluctuation range is analyzed based on their corresponding characterization index called the variance. In Section 4, the law of the influence of spatiotemporal diurnal modulation of atmospheric turbulence on the random fluctuation rate of speed is analyzed by variation index. Furthermore, this method is extended to analyze wind power fluctuation rate. In Section 5, the influence of wind speed and wind power intermittency on turbulent spatiotemporal daily modulation is analyzed based on the ramp duty ratio index. Moreover, another wind power intermittency characterization parameter without daily modulation is analyzed by the start/stop frequency of the wind turbines. In Section 6, considering the fluctuation range of wind speed and wind power as examples, the quantitative characterization modeling of the uncertainty and its introduction strategy in the evaluation of real-time frequency modulation capability of power grid are used to explore the feasibility of such spatiotemporal daily modulation characteristics. The seventh section is the conclusion and prospection of the paper.

2. Intrinsic Properties of Spatiotemporal Diurnal Modulation in Atmospheric Turbulence

2.1. Diurnal Modulation and Spatiotemporal Characteristics of Atmospheric Turbulence. Atmospheric turbulence is an intrinsic factor of uncertainty in wind speed and the power output of wind turbine. This uncertainty performance in atmospheric motion is attributed to the random fluctuation of various sizes superimposed by its average wind speed and wind direction. Similarly, it also applies to the power output of wind turbines.

Turbulence is a different kind of motion that takes place in the atmospheric boundary layer (ABL). The factors affecting its formation differ during day and night with apparent diurnal periodic characteristics. Generally, the turbine hub height is in the ABL. The diurnal variation of the sun rising in the east and setting in the west highly

influences the convective motion of the ABL in a flat terrain where surface roughness is uniform. Therefore, diurnal variation can reflect the change in an underlying flat surface in the boundary layer. However, the turbulence intensity in the underlying surface layer of a complex mountainous area is affected by the vortex flow around the local landform, leading to the formation of intense turbulence as well, which is stronger than that caused by sunlight, where diurnal period characteristics are submerged. This influence, especially, is greater when the altitude closer to the ground, leading to a greater turbulence than that at a higher altitude. The diurnal periods of flat areas, such as plains or plateaus, have common features. For example, the time of maximum turbulence intensity is similar. Therefore, the diurnal period has spatiotemporal characteristics, which are derived from the intrinsic characteristics of the turbulence in ABL.

The process of controlling one parameter of a signal with another signal is defined as modulation in communication systems. Based on the influence of the diurnal period on turbulence and wind speed, this diurnal period can be considered as the modulation of wind speed. In other words, the diurnal period is a diurnal modulation process of wind speed caused by a longer weather process, and a diurnal modulation process of the atmosphere is the physical mechanism for the existence of diurnal period phenomenon in the turbulent part of wind speed and wind power. Therefore, several characterization indexes of wind speed uncertainty are affected by spatiotemporal diurnal modulation of atmospheric turbulence, such as the average wind speed of wind farm, the output power of the wind turbines, the random fluctuation part and internal intermittency, and the average temperature (heat flux) of the wind field. However, the descriptive parameters that cannot reflect the inherent characteristics of wind speed are not affected.

2.2. Other Recommendations. The scheduling, control, and planning arrangements in a power system have different time scales. The day-ahead prescheduling planning takes one day as a computation period, while the real-time scheduling and optimization control takes hours as the time scale. These correspond to the 0~24 h day-ahead forecast and 0~4 h ultra-short-term forecast of wind power, respectively. Therefore, if the parameters related to wind power prediction, such as the mean time value of wind speed and wind power, wind power uncertainty, and intermittency, especially for the hourly scale statistic rules, have diurnal modulation characteristics, it will have a guiding significance for the day-ahead prescheduling planning and real-time scheduling of primary and secondary frequency modulation in the new energy power system after large-scale wind power grid connection.

Autocorrelation analysis is a mathematical tool used to find repeated patterns and analyze value functions or sequences of signal processing, such as a periodic signal masked by noise. As for periodic sequences, the autocorrelation sequence shows periodic change [43].

3. Influence Law of Turbulence Diurnal Modulation on Wind Power Random Fluctuation Range

3.1. Quantitative Characterization of Wind Power Fluctuation Range. Generally, the actual wind speed can be divided into hourly average wind speed and turbulent wind speed based on Reynolds averaging. Research results show that the turbulent part of the wind speed depends on the average time strongly. A turbulence intensity model in international IEC standard realizes the wind speed fluctuation range (intensity). Subsequently, a universal model of turbulence intensity was proposed [43], as given by the following equation:

$$TI = \frac{\sigma}{\overline{v}} = \alpha \cdot \overline{v}^{-\beta} + c, \tag{1}$$

where σ is variance of wind speed turbulence; $\overline{v}$ is the mean wind speed per hour; α, β, and c are constants. Similarly, the model of single power and relative variance for wind turbine or wind farm can be established. I_p, the wind power fluctuation intensity, is the characteristic parameter of power fluctuation range [40]. The relative variance is the unit value of the residual fluctuation components after subtracting the mean value from the actual wind power. Based on the wavelet algorithm, the multiscale fluctuation intensity suitable for frequency modulation capability evaluation can be obtained as the minute-scale wind power fluctuation intensity I_{pm} and secondary wind power fluctuation intensity I_{ps}, shown in equations (2) and (3), respectively.

$$I_{pm} = \frac{\sigma_m}{\overline{P}} = \alpha_m \times \overline{P}^{-\beta_m} + c_m, \tag{2}$$

$$I_{ps} = \frac{\sigma_s}{\overline{P}} = \alpha_s \times \overline{P}^{-\beta_s} + c_s, \tag{3}$$

where σ_m is the instantaneous standard deviation of wind power in minute-scale; σ_s is the instantaneous standard deviation of secondary wind power; $\overline{P}$ is the mean wind speed per hour; α_m, β_m and c_m are the fitting constants of minute-scale wind power fluctuation intensity; α_s, β_s, and c_s are the fitting constants of secondary wind power fluctuation intensity.

3.2. The Influence Law on Wind Speed Fluctuation Range. There are many factors that affect the actual wind speed fluctuation uncertainty, and the three-parameter power law model also has a fitting error under certain conditions. The fitting error is defined as follows:

$$e = \sigma - \left(\alpha \cdot \overline{v}^{-\beta} + c\right)\overline{v}, \tag{4}$$

where σ is the residual fluctuation standard deviation of the actual wind speed after eliminating average.

Figure 1 shows the autocorrelation analysis of the fitting error. It can be seen that the fitting error has strong diurnal period characteristics, and the diurnal period change pattern is different for the four seasons.
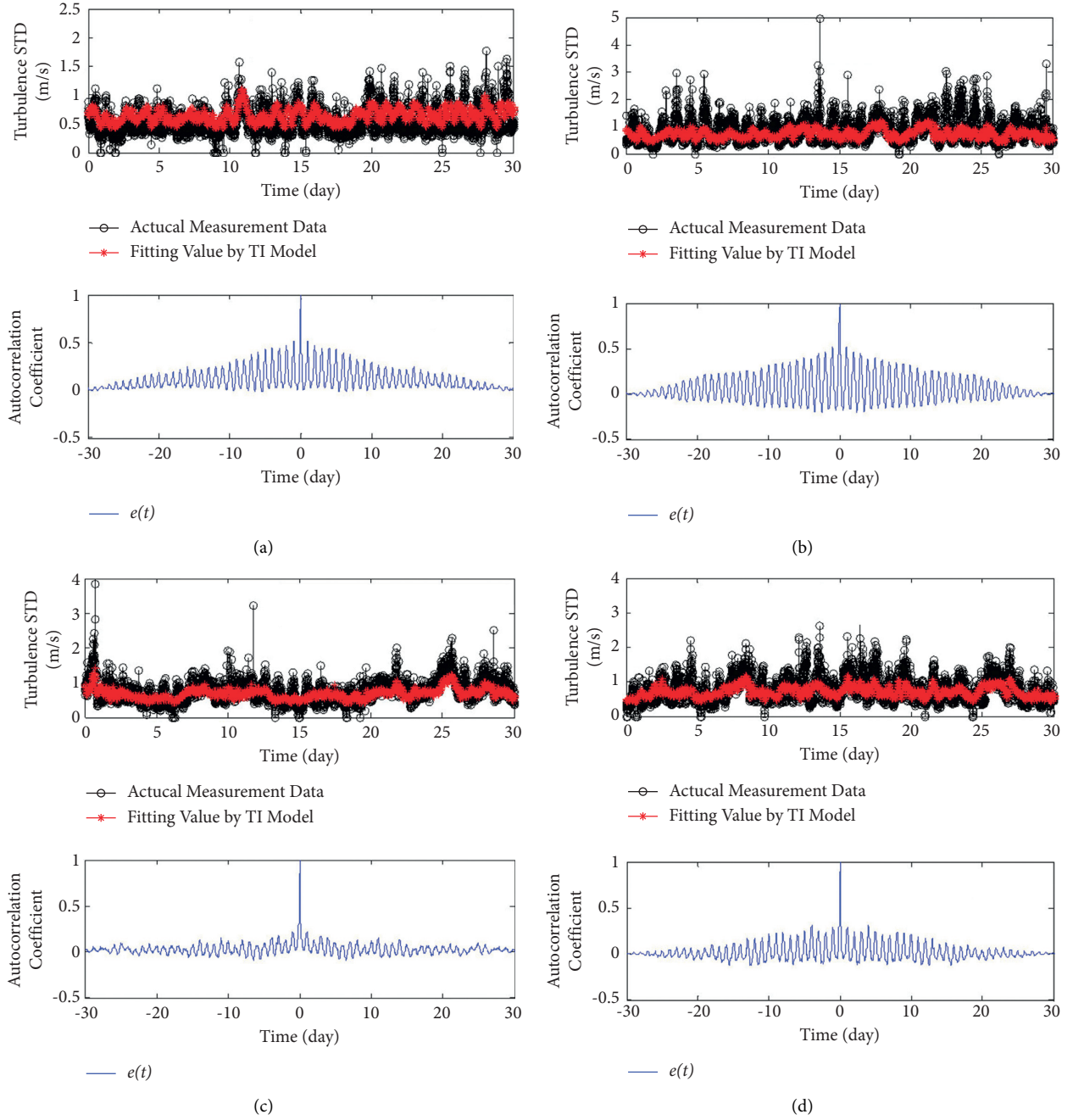
FIGURE 1: The fitting error of the model and its diurnal period pattern. (a) Wind data in spring. (b) Wind data in summer. (c) Wind data in autumn. (d) Wind data in winter.

There is an obvious diurnal periodic change pattern of turbulence intensity combining the above analysis. Figure 2 shows the variation law of the model parameters within 24 h obtained by fitting the data in Figure 1. It indicates that the model parameters are variable throughout the daily cycle, and the differences are relatively large, which cannot be ignored.

Viewing from the perspective of wind turbine scheduling and controlling, the average turbulence intensity is defined by the effective section in the middle of the cut-in/cut-out speed of the universal fan. As shown in

Figure 3, the average turbulence intensity in the wind speed section, ranging from 3 m/s to 25 m/s, corresponds to the hub height of the fan. The calculation used is shown in the following equation:

$$\bar{I} = \frac{1}{n} \sum_{i=1}^{n} I_i, \tag{5}$$

where I_i is the turbulence intensity corresponding to the average speed range from 3 m/s to 25 m/s; n is the number of samples within the average speed range from 3 m/s to 25 m/s.
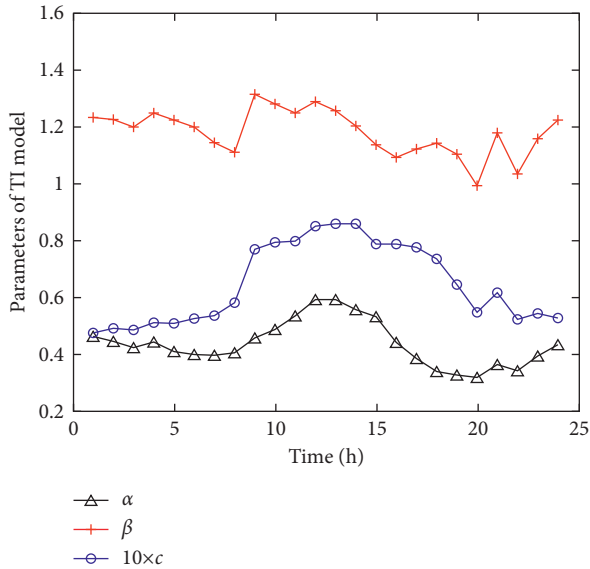
FIGURE 2: Model parameter variation within 24 h.
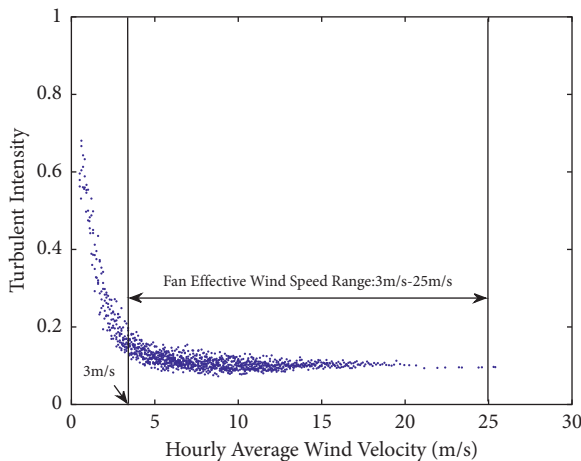
FIGURE 3: The hub height's average turbulence intensity.

Figure 4 shows the diurnal period variation pattern of different wind speed turbulence intensities obtained by the fitting model. Figure 5 shows the diurnal period patterns of different wind speed turbulence intensities obtained from the measured data. All of them prove the existence of the diurnal period change pattern of turbulence intensity.

Figures 6(a)–6(c) show seasonal and monthly variation patterns of turbulence intensity in diurnal period, including diurnal variation patterns of turbulence intensity in four seasons and twelve months. It can be seen that the diurnal period pattern of turbulence intensity in this region is affected by the season and month. Generally, the average turbulence intensity is the highest in spring and the lowest in winter. The maximum value of average turbulence intensity in a single month occurs in May and September. In terms of the 24 h variation in a day, the maximum value occurs around 14 : 00~15 : 00, and the minimum value occurs around 20 : 00 of the same day~08 : 00 of the next day. The maximum value occurs earlier in winter compared to other seasons.

All the investigated wind farms were located in typical areas of northern China, but the diurnal cycle patterns of the actual offshore wind farms were different. Land turbulence intensity whose driving energy is predominantly from the sun has a diurnal cycle variation pattern, and the difference is mainly due to the difference in light intensity and illumination time during different seasons and months. It means that the diurnal modulation of atmospheric turbulence has spatiotemporal characteristics.

To further illustrate the issue, the one-year data from three wind farms in different regions are analyzed based on effective average turbulence intensity parameters. Figure 6(d) shows the diurnal characteristics of average turbulence intensity in different regions. The diurnal cycle patterns of three wind farms differ with the time and location. The main reason for this is the wind farms located in the different latitudes and longitudes. It means that the time of sunrise and sunset is different, resulting in a time deviation of the maximum value.

As shown in Figure 7, the influence of the sunshine on the diurnal variation is weakened by complex topography and varied surface roughness. Figure 8 shows the maximum value of the average turbulence intensity corresponding to the fan at the lowest altitude is nearly twice that of its maximum value corresponding to the fan at the highest altitude. Therefore, complex topography has a great influence on the average turbulence intensity. The lower the altitude, the greater it is affected by topography and roughness. However, literature [43] does not elaborate on these aspects.

3.3. The Influence Law of Turbulence Diurnal Modulation on Wind Power Fluctuation Range. The diurnal cycle characteristics of wind power fluctuation intensity are not mentioned in the existing literature. Figure 9 shows the diurnal cycle characteristics of wind power fluctuation intensity of the different fans in the same wind field. Due to the variance of wind speed has diurnal cycle characteristics, the variance of wind power also has diurnal cycle characteristics. Moreover, the wind power fluctuation intensity corresponding to different power values also has diurnal cycle characteristics. However, after the conversion of wind speed into energy by means of the fan, the regularity of the diurnal cycle change of wind power fluctuation intensity is not as apparent as that of the diurnal cycle change of wind speed turbulence intensity.

4. Influence Law of Turbulence Diurnal Modulation on Wind Power Random Fluctuation Rate

4.1. Quantitative Characterization of Wind Power Fluctuation Rate. Variational analysis is introduced in [41] to establish a computing model of instantaneous wind speed variation based on wavelet algorithm. In addition, the quantitative characterization of wind speed fluctuation rate is realized based on the dependence of wind speed variation on hourly average wind speed, similar to turbulence intensity. The
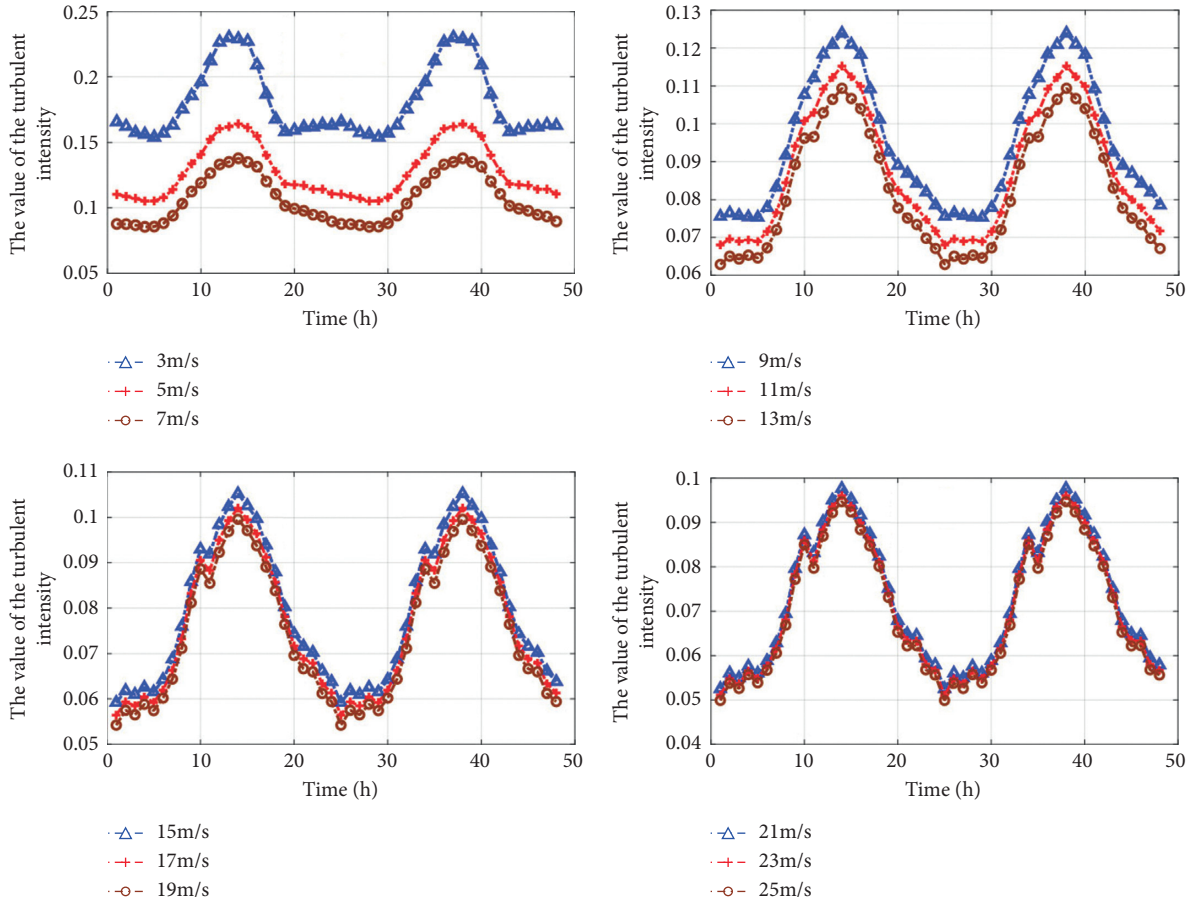
FIGURE 4: The variation of turbulence intensities for different wind speeds within 24 h using fitting model.
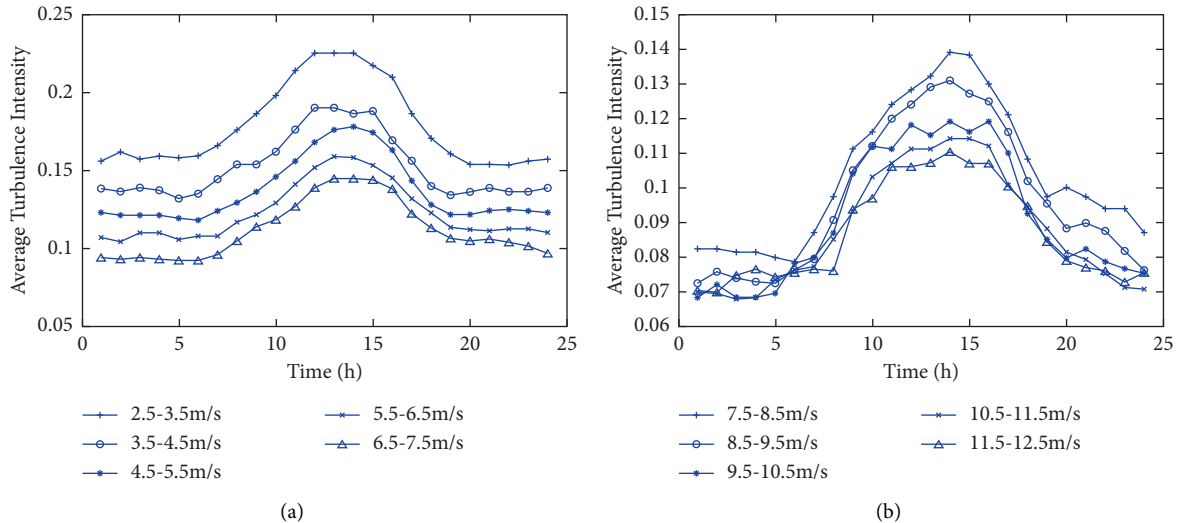
FIGURE 5: The variation of the mean turbulence intensities value within 24 h time interval (actual data). (a) Wind speed range from 2.5 m/s to 7.5 m/s. (b) Wind speed range from 7.5 m/s to 12.5 m/s.

variation of wind power also has the same characteristics. We further defined the variable intensity of wind power χ_p as the derived parameter to describe wind power variation quantitatively, as given by the following equation:

$$\chi_p = \frac{\left[\gamma_p^*(\Delta t) \right]^{1/2}}{\overline{p}^*} = \alpha \cdot \left(\overline{p}^* \right)^{-\beta} + c, \tag{6}$$
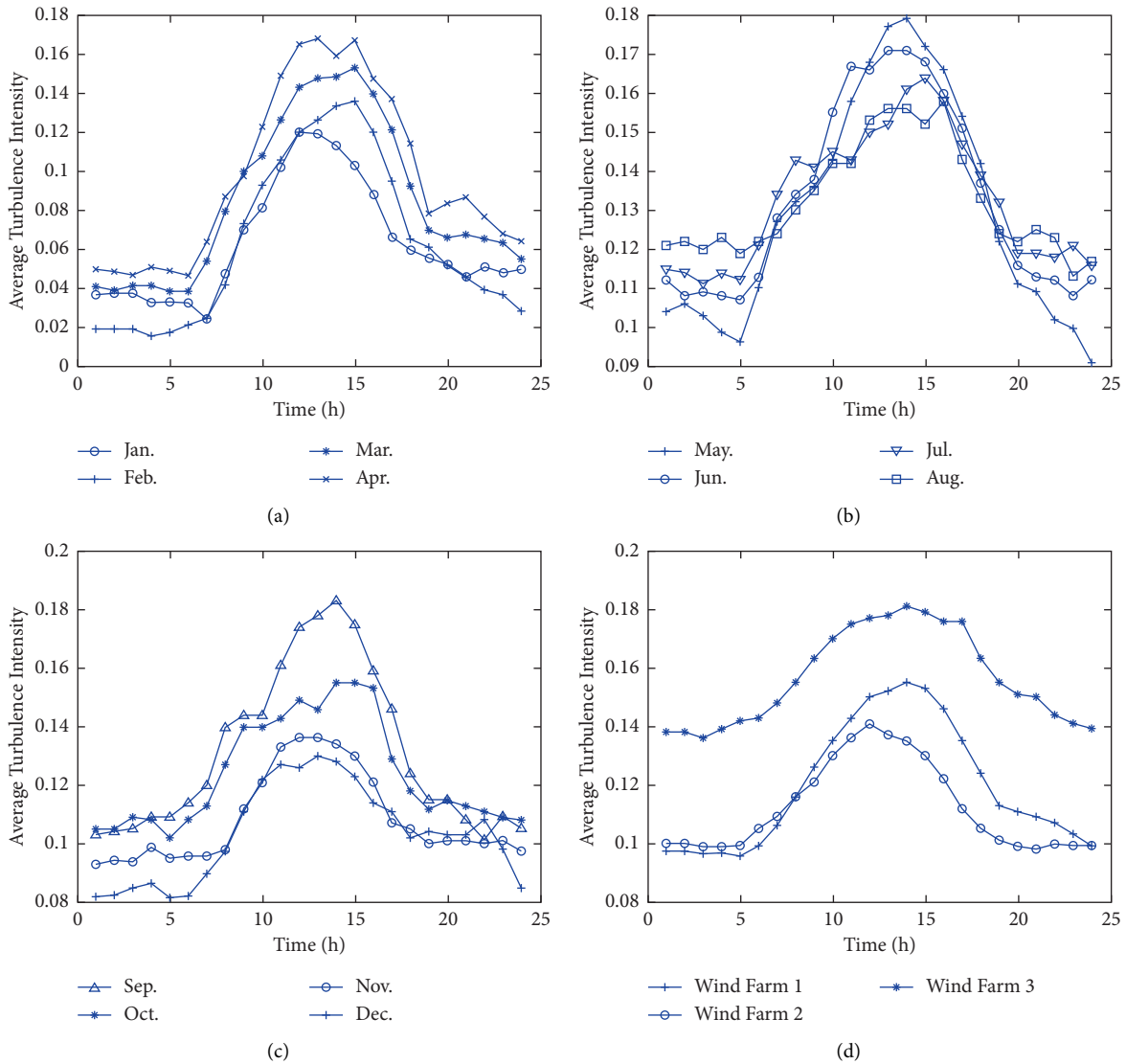
FIGURE 6: The diurnal cycle variation of TI during different seasons and months of a year. (a) The diurnal cycle from January to April. (b) The diurnal cycle during May and August. (c) The diurnal cycle from September to December. (d) The diurnal cycle variation of mean TI in three different farms.
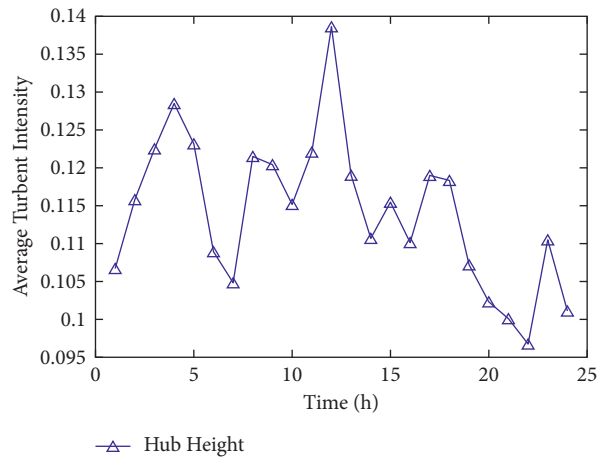
FIGURE 7: The diurnal cycle variation of TI in land farm with complex terrain.
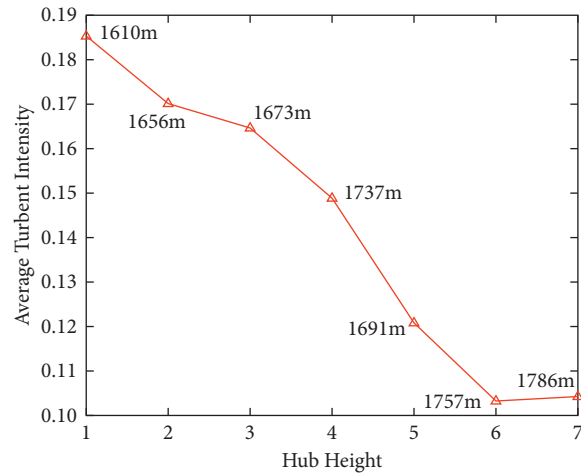
FIGURE 8: The diurnal cycle variation of TI in land farm with complex terrain.
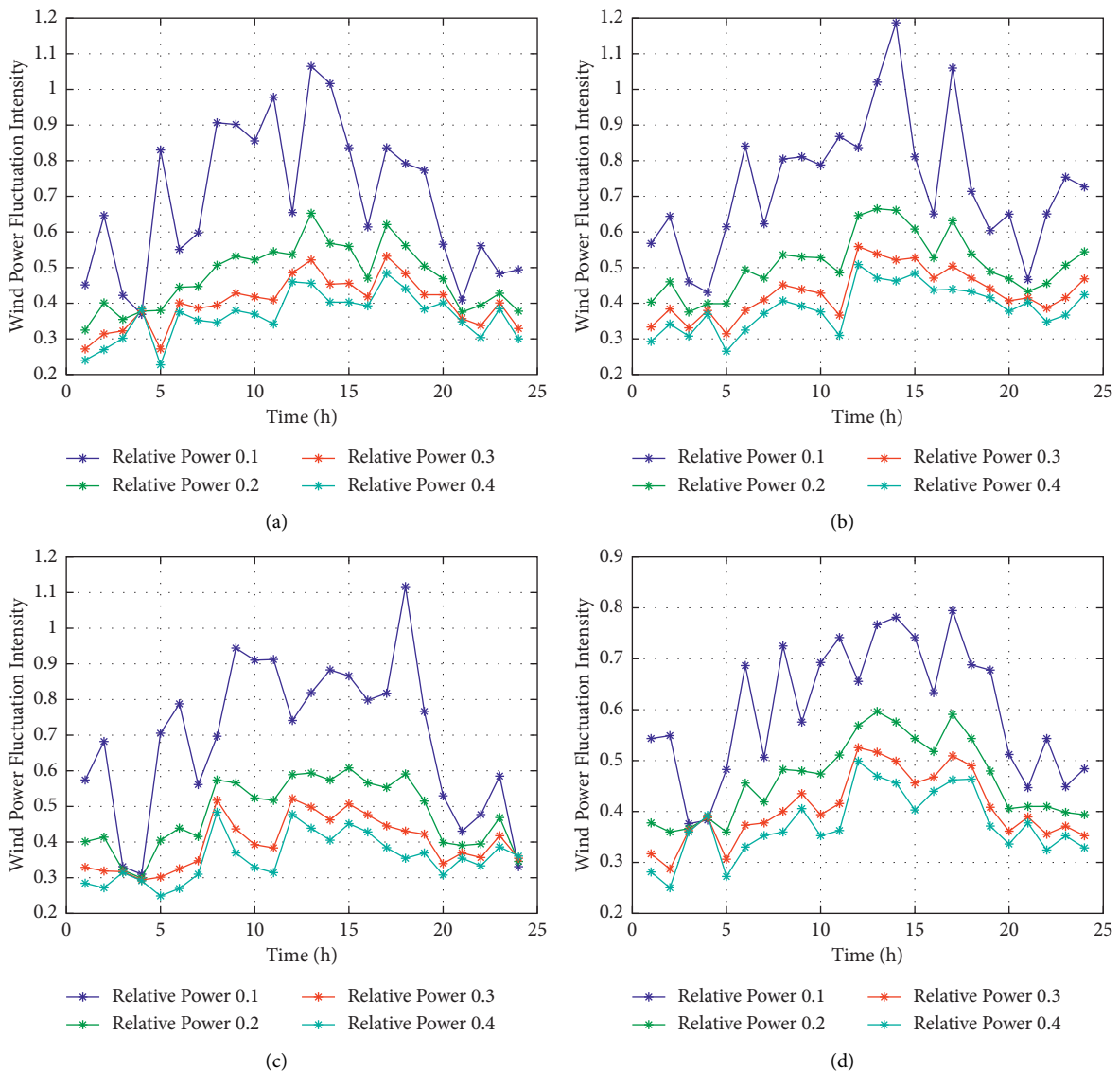
FIGURE 9: The diurnal cycle of wind power variance in same wind field. (a) #19 turbine. (b) #39 turbine. (c) #69 turbine. (d) #119 turbine.

where $\gamma_p^*(\Delta t)$ is the relative variation of wind power; $\overline{p}^*$ is the unit value of the average hourly power; α, β, and c are fitting constants.

4.2. Influence Law of Turbulence Diurnal Modulation on Wind Speed Fluctuation Rate. In [41], one-month time series data was used for autocorrelation analysis, which adopted multitime interval variation calculation. The results show that the time lag of the variation function is a periodic component in the time series, and the variation period is 1 day. The regularity of diurnal cycle change pattern remains constant with an increase in time interval. The diurnal cycle regularity becomes weaker only when the time interval reaches a certain level. The variation function of wind speed has the same diurnal cycle pattern as the variance. The time series $\gamma(\Delta t)$, as in [41], has different time lags, and the diurnal pattern of wind speed variation function is analyzed in detail. $\gamma(\Delta t)$ is large from 8 a.m. to 6 p.m., especially at its peak between 12 a.m and 4 p.m, and relatively small during other times of the day. On the basis of [41], further analysis reveals that diurnal cycle characteristics are not related to the speed, and speeds at different sizes have similar daily cycle characteristics shown in Figure 10.

4.3. The Diurnal Period Characteristics of Wind Power Variogram. Based on [41], the universality of the diurnal period characteristics of wind power variogram is further analyzed. As shown in Figure 11, the measured wind power data of multiple Wind generators are selected. It shows that wind power variogram is modulated by the diurnal process and has an obvious diurnal period characteristics. However, none of the present literature mentions the diurnal cycle characteristics of wind power variogram. Therefore, we extend the results of our research from characteristics of wind speed to its producing wind power. In power grid dispatching, the wind speed during noon has strong fluctuations at a fast rate. Therefore, when the disturbance caused by the power stroke of the grid is suppressed, it is necessary to reasonably configure fans with different suppression capacities to ensure safe and stable operation of the grid.

5. Influence Law of Turbulence Diurnal Modulation on Wind Power Intermittency

In [45], a quantitative description method of intermittency is proposed based on the duty ratio of abrupt change in wind speed, which considers the atmospheric turbulence physical essence. Furthermore, in [46], this kind of quantitative description method is extended to the definition of wind power intermittency, which is of great significance to the power system.

In [45, 46], similar autocorrelation has been adopted to analyze the measurement parameters of wind speed and its generating power intermittently in detail. It was found that the abrupt duty cycle parameters of wind power had obvious diurnal cycle characteristics. Similar to the diurnal cycle of wind power variance and variogram parameters, the steep

duty ratio between 8 a.m. and 8 p.m. is larger than other times. It reaches its daily peak between 12 a.m and 4 p.m. With similar statistical analyze methods, the results calculated on sufficient data from this farm but other months and other wind farms are identical. A study on the daily-cycle characteristic of the abrupt duty cycle of wind speed at different time intervals showed that the diurnal cycle phenomenon of the abrupt duty cycle of wind speed still exists. This proves the universality of the diurnal cycle phenomenon. The abrupt change in the diurnal cycle phenomenon of wind speed in duty cycle indicates that the wind power intermittency is stronger during the day than during night. Wind power intermittency reaches its daily peak between 12 a.m and 4 p.m. Therefore, the intermittent characterization index reflects the intrinsic characteristics of wind power uncertainty.

Although several parameters are used to characterize wind power intermittency, these parameters that cannot describe the nature of the intermittent wind speed and atmospheric turbulence also have no modulation effect.

6. Applications

6.1. Improved Wind Power Uncertainty Model Considering the Diurnal Modulation Characteristics of Atmospheric Turbulence. The study of the diurnal modulation characteristics of atmospheric turbulence on wind power uncertainty has the following practical significances for the utilization of wind power:

(1) The diurnal period characteristics indicate that some factors of turbulence will affect the parameters of wind power variance and variogram model. Therefore, the region, topography, seasons, and weather patterns (such as sunny and rainy weather) will influence atmospheric turbulence.

(2) The diurnal period of wind speed variance is a reflection of the random fluctuation of wind speed. Therefore, wind speed fluctuation range should be larger when turbulence intensity is higher, instead of a fluctuation range of equal widths.

(3) According to the diurnal period of the wind speed variogram, the wind power change rate is unstable at different times during a day, which means that system operators must change their strategies in different states. More specifically, complementary power and control strategies that respond faster must be provided to ensure the safe and efficient operation of the new energy power system when the light intensity is stronger.

(4) In addition, the diurnal cycle characteristics of wind speed variance and variogram also have reference value of the electricity price bidding for wind farm. The diurnal cycle characteristics of intermittency indexes have a similar significance for the operation of wind farms [53].

The diurnal cycle characteristics of wind power have a great impact on the precision of qualitative model fitting.

(a)

(b)
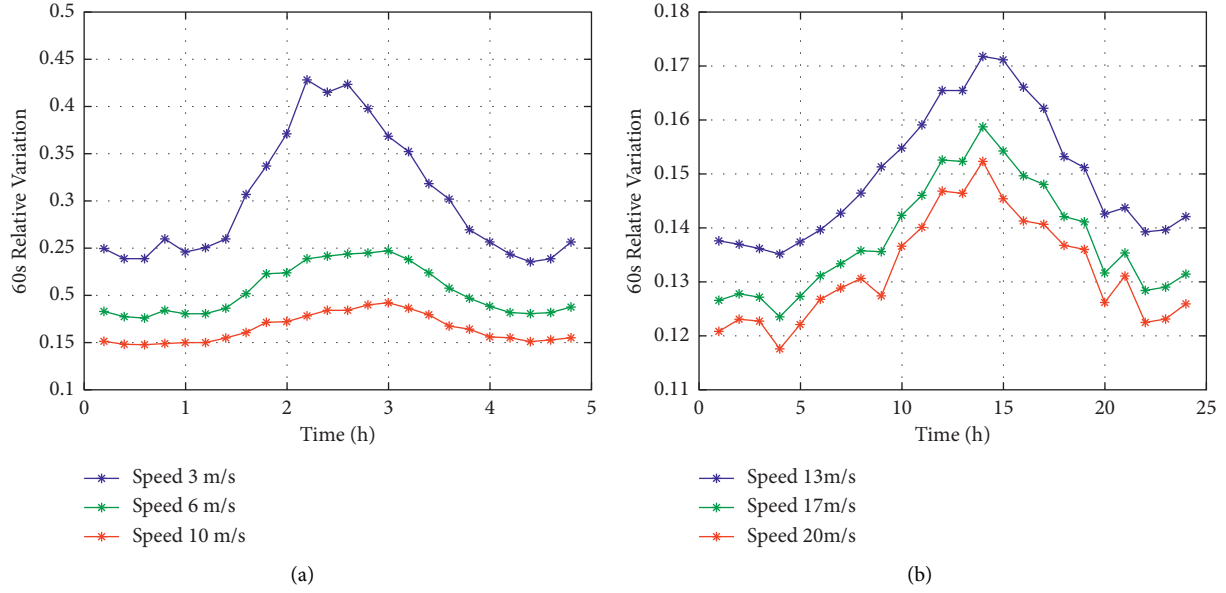
FIGURE 10: The diurnal cycle of different wind speed variation. (a) Wind speed: 3 m/s, 6 m/s, 10 m/s. (b) Wind speed: 13 m/s, 17 m/s, 20 m/s.

According to the fitting effect of time-sharing model in [43], we can further optimize the variation and intermittency models. Taking the variance model as an example, the model considering time-varying parameters has the higher fitting accuracy [54]. The time-varying parameter model is shown in the following equation:

$$\frac{\sigma(t)}{\overline{v}(t)} = \alpha(t) \cdot \overline{v}(t)^{-\beta(t)} + c(t), \tag{7}$$

where $\sigma(t)$ is the variance of actual measurement wind speed; $\overline{v}(t)$ is the mean actual wind speed per hour; $\alpha(t)$, $\beta(t)$, and $c(t)$ are a group of constants obtained by data fitting and time-sharing modeling.

As shown in Figure 12, the fitting error of this model considers the characteristics of diurnal cycle. Compared with Figure 1, the results show that this model has a great fitting accuracy rather than the original one (equation (1)).

The following four indicators of quantitative fitting effect are statistically analyzed.

$$\text{MNRE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(y_{ri} - y_{fi})}{y_{ri}} \right|,$$

$$\text{MNSE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(y_{ri} - y_{fi})^2}{y_{ri}^2} \right|,$$

$$\text{HASL} = k_{(\rho_k > 0.6)}, \tag{8}$$

$$\text{MAC} = \frac{1}{n} \sum_{k=1}^{n} |\rho_k - \min(\rho_k)|,$$

where MNRE and MNSE are mean normalized relative error and mean normalized square error, respectively. y_{ri} and y_{fi} are the actual variance values calculated by using wind speed

data and obtained by using model fitting, respectively. HASL and MAC are the mean of the highly autocorrelated step size and autocorrelation coefficient, respectively. ρ_k is the autocorrelation coefficient. $k_{(\rho_k > 0.6)}$ is the step size corresponding to $\rho_k > 0.6$.

Table 1 shows a comparison of the results. It can be seen that the fitting-effect of the time-varying parameter model is better than that of the original model.

6.2. Example Analysis of Frequency Modulation Capability Estimation Considering the Diurnal Characteristic of Wind Power Fluctuation Intensity.

According to [40], the fluctuation range measurement of wind power can be analyzed combined with the primary frequency modulation (PFM) and secondary frequency modulation (SFM) characteristics of thermal power units. The units for frequency modulation can be rationally configured to meet the analysis requirements of frequency modulation. The fluctuation intensity of minute and second wind power in a wind farm during a day was fitted, and the corresponding fluctuation time series were calculated as follows:

(1) When curve fitting is carried out at all time periods, the corresponding single-fitting parameter in equation (2) can be obtained. Based on the fitting model, wind power fluctuations in different time scales are estimated. The allocation scheme 1 is as follows: the proportion of primary FM unit is 80%. The inequality of turbine generator unit $\delta iA = 0.05$. The proportion of secondary FM unit is 40%, and the integral gain KA of secondary FM channel is 0.25. At this time, formula 20 in reference [40] (DFPRA) of the primary frequency modulation capability of the system increases from 7.9523 to 15.9045, and formula 21 in reference [40] (DSPR) of the secondary frequency modulation capability increases from 11.9636 to 23.9273.
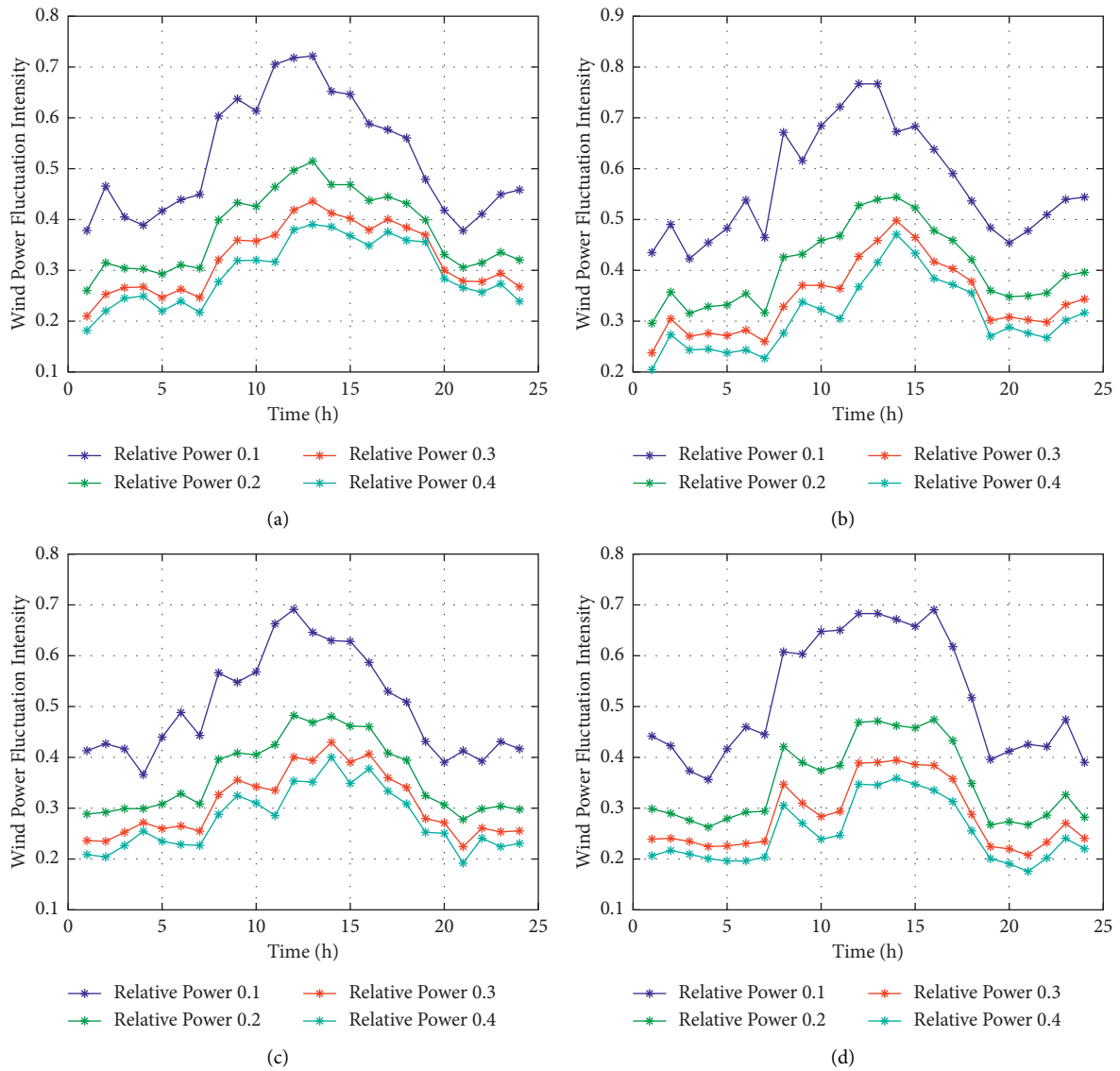
Figure 11: The diurnal cycle of different wind turbine power variogram in wind field. (a) #19 turbine. (b) #39 turbine. (c) #69 turbine. (d) #119 turbine.
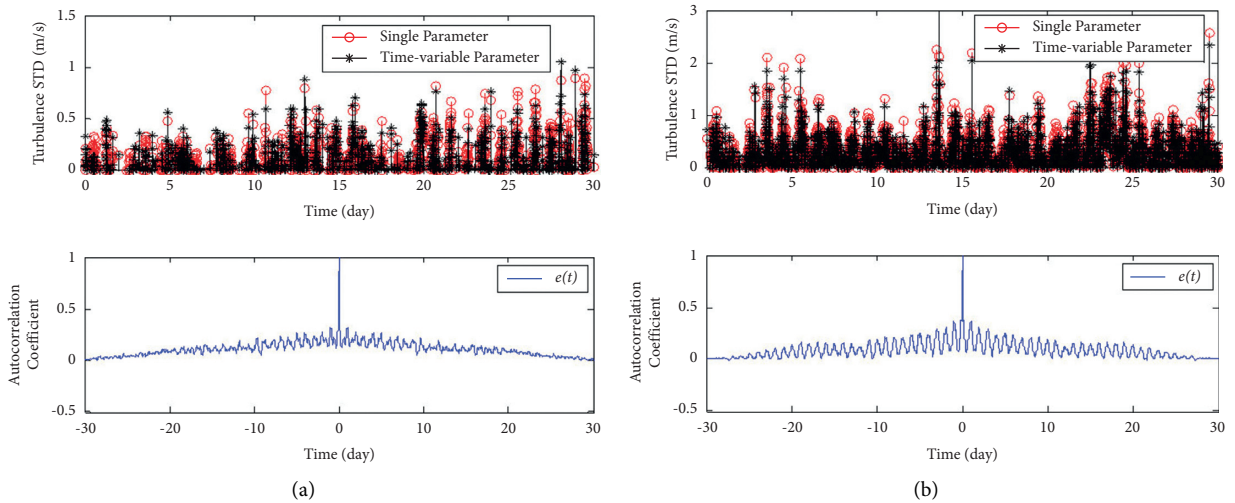
Figure 12: Continued.
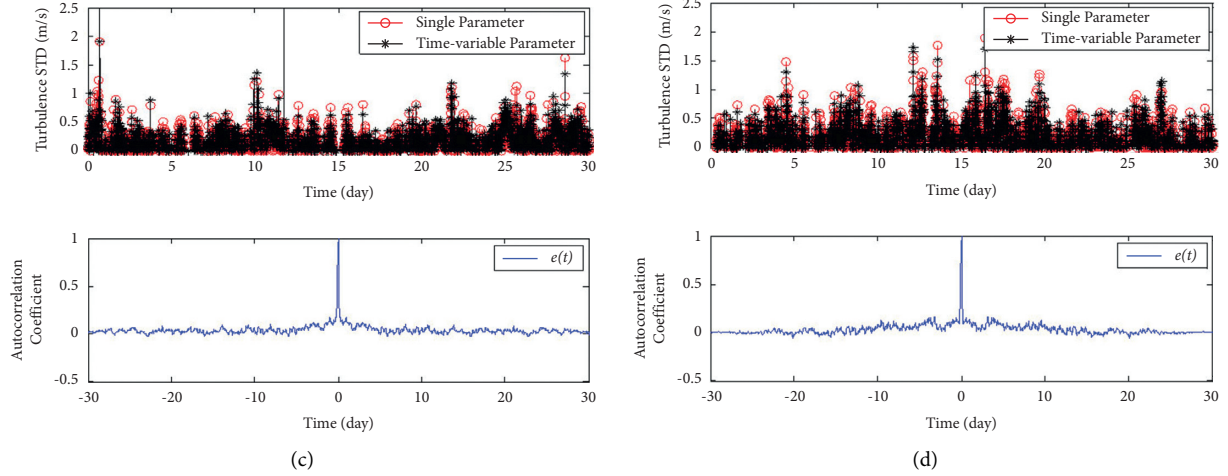
FIGURE 12: The fitting error of time variant model considering daily cycle model. (a) Wind data in spring. (b) Wind data in summer. (c) Wind data in autumn. (d) Wind data in winter.

TABLE 1: Statistical analysis of the fitting error.

Dataset	Original model				Time-varied model			
	MNRE(%)	MNSE(%)	HASL	MAC	MNRE(%)	MNSE(%)	HASL	MAC
1	47.53	13.90	8	0.1222	38.78	9.56	3	0.1091
2	36.12	15.21	8	0.2661	28.40	10.44	2	0.0943
3	27.73	7.68	2	0.1181	25.11	6.73	1	0.0658
4	33.41	10.87	5	0.1493	27.71	7.88	2	0.0853
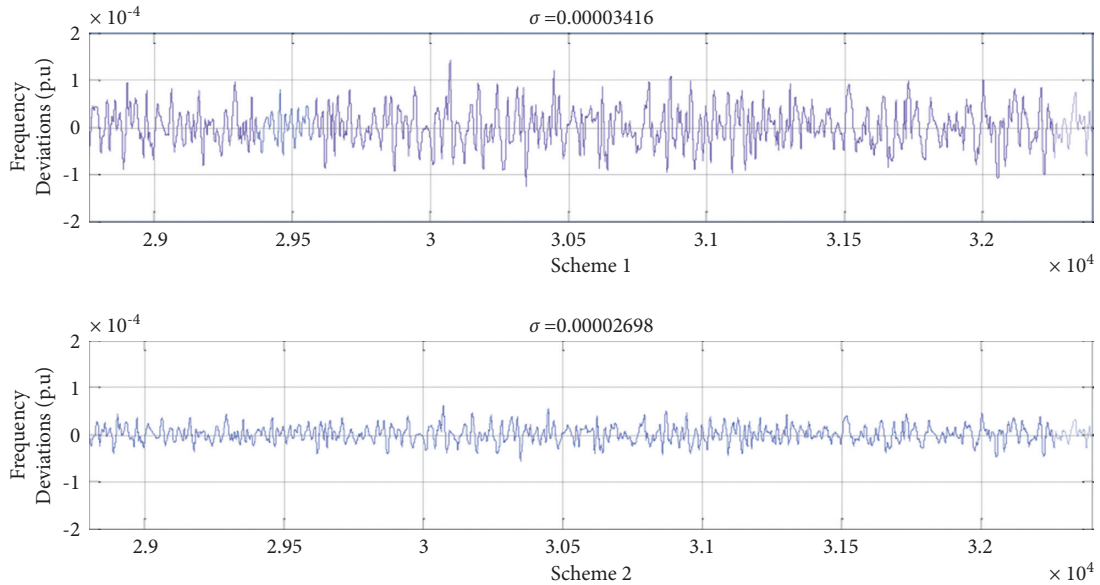
FIGURE 13: Comparison of simulation results of two different schemes.

(2) According to the diurnal modulation characteristics, the data are fitted during a day. The corresponding power law model of wind power fluctuation in this period is obtained subsequently. Based on the law model of time-sharing wind power fluctuation, the wind power fluctuation during this period can be estimated accurately. So, scheme 2 is as follows: the primary FM unit of share A is increased from 40% to 50%. The range of turbo-generator set rate δiA is reduced from 0.05 to 0.025. The secondary FM unit share is increased from 20% to 40%. The integrator gain KA is increased from 0.25 to 0.5. Such a system

has the primary frequency capacity DFPRA increasing from 7.9523 to 19.9004, and secondary frequency modulation capacity DSPRA increased from 11.9636 to 60.0533. The system's frequency deviation after adopting the two schemes is shown in Figure 13. It can be seen that the standard deviation of the frequency fluctuation of the system $\sigma = 0.00003416$ when scheme 1 is used in region A being reduced to $\sigma = 0.00002698$ when scheme 2 is used. Therefore, the time-sharing modeling method, considering the daily modulation characteristics, can achieve accurate estimation of wind power fluctuations, realize the optimal allocation of FM units, and provide a reference for the power grid dispatchers to rationally allocate FM units.

7. Conclusion

Considering the physical nature of atmospheric motion, the wind power uncertainty is studied in the present research. The influence of the diurnal cycle characteristic of atmospheric turbulence on the fluctuation of wind speed and wind power is studied in this paper. The results show that this characteristic itself is affected by season, latitude, longitude, and topography. However, its influence on wind power fluctuation is weakened after the energy conversion process. Then, we analyze the influence of the diurnal cycle characteristic on the random fluctuation rate of wind power, and wind power intermittency. The results show that this parameter has no influence on the wind speed and the proportion of wind speed steepness but affects the duty ratio of abrupt change in wind speed.

Based on this research, we propose a time-varying model considering the diurnal cycle characteristic, and the fitting error of wind speed and its power generation fluctuation models are all smaller compared with the original models. At the same time, the wind power fluctuation is estimated based on the law model of time-sharing wind power fluctuation. In the two simulation schemes, the standard deviation of the system frequency fluctuation is reduced to 0.00003416 and 0.00002698, respectively, which verifies the validity of the model.

This paper presented the spatiotemporal diurnal modulation characteristics of atmospheric turbulence from the perspective of real-time scheduling and optimal control of power systems. This is of practical significance for power grid dispatchers to adjust the frequency modulation resources in real time and ensure the safety and reliability of system frequency.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

A preprint of the article has been published (DOI: https://doi.org/10.21203/rs.3.rs-1051667/v1).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] P. Scarabaggio, S. Grammatico, R. Carli, and M. Dotoli, "Distributed demand side management with stochastic wind power forecasting," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 1, pp. 97–112, 2022.

[2] H. Zhang, X. Hu, and H. Cheng, "Coordinated scheduling of generators and tie lines in multi-area power systems under wind energy uncertainty," *Energy*, vol. 222, Article ID 119929, 2021.

[3] C. Wu, K. Luo, Q. Wang, and J. Fan, "A refined wind farm parameterization for the weather research and forecasting model," *Applied Energy*, vol. 306, Article ID 118082, 2022.

[4] H. H. Tsao, Y. G. Leu, and L. F. Chou, "A center-of-concentrated-based prediction interval for wind power forecasting," *Energy*, vol. 237, Article ID 121467, 2021.

[5] X. Lyu, Y. Jia, T. Liu, and S. Chai, "System-oriented power regulation scheme for wind farms: the quest for uncertainty management," *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 4259–4269, 2021.

[6] P. Lu, L. Ye, Y. Zhao, B. Dai, M. Pei, and Y. Tang, "Review of meta-heuristic algorithms for wind power prediction: methodologies, applications and challenges," *Applied Energy*, vol. 301, Article ID 117446, 2021.

[7] A. A. Juárez, A. M. Araújo, and J. S. Rohatgi, "Development of the wind power in Brazil: Political, social and technical issues," *Renewable and Sustainable Energy Reviews*, vol. 39, no. 6, pp. 828–834, 2014.

[8] Z. Y. Zhao, P. H. Wu, and B. Xia, "Development route of the wind power industry in China," *Renewable and Sustainable Energy Reviews*, vol. 34, no. 3, pp. 1–7, 2014.

[9] J. Shen, X. Song, Z. Ming, W. Yi, W. Yuejin, and L. Xiaoli, "Low-carbon development strategies for the top five power generation groups during China's 12th Five-Year Plan period," *Renewable and Sustainable Energy Reviews*, vol. 34, no. 34, pp. 350–360, 2014.

[10] X. Baichen, S. Lifeng, L. Wenhua, and Z.-Y. Lu, "Study on China's wind power development path Based on the target for 2030," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 197–208, 2015.

[11] J. G. Slootweg and W. L. Kling, "The impact of large scale wind power generation on power system oscillations," *Electric Power Systems Research*, vol. 67, no. 1, pp. 9–20, 2003.

[12] Y. V. Makarov, C. Loutan, and J. Ma, "Operational impacts of wind generation on California power systems," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1039–1050, 2009.

[13] G. Callegari, P. Capurso, F. Lanzi, M. Merlo, and R. Zaottini, "Wind power generation impact on the frequency regulation: study on a national scale power system," in *Proceedings of the 14th International Conference on Harmonics and Quality of*

Power-ICHQP 2010 Wind Energy, pp. 1–6, Bergamo, Italy, September 2010.

[14] G. Ren, J. Liu, J. Wan, and Y. Guo, "Overview of wind power intermittency: impacts, measurements, and mitigation solutions," *Applied Energy*, vol. 204, pp. 47–65, 2017.

[15] J. Wu, Z. X. Wang, and G. Q. Wang, "The key technologies and development of offshore wind farm in China" Renewable and Sustainable," *Energy Reviews*, vol. 34, pp. 453–462, 2014.

[16] Y. He, Y. Pang, J. Zhang, T. Xia, and T. Zhang, "Feed-in tariff mechanisms for large-scale wind power in China," *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 9–17, 2015.

[17] W. Wang, D. Zeng, J. Liu, Y. Niu, and C. Cui, "Feasibility analysis of changing turbine load in power plants using continuous condenser pressure adjustment," *Energy*, vol. 64, pp. 533–540, 2014.

[18] W. Wang, S. Jing, Y. Sun, Y. Niu, D. Zeng, and C. Cui, "Combined heat and power control considering thermal inertia of district heating network for flexible electric power regulation," *Energy*, vol. 169, pp. 988–999, 2019.

[19] A. M. Howlader, N. Urasaki, A. Yona, T. Senjyu, and A. Y. Saber, "A review of output power smoothing methods for wind energy conversion systems," *Renewable and Sustainable Energy Reviews*, vol. 26, no. 10, pp. 135–146, 2013.

[20] O. P. Mahela, A. G. Shaik, and L. Kazmerski, "Comprehensive overview of grid interfaced wind energy generation systems," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 260–281, 2016.

[21] J. Wang, Y. Song, L. Feng, and R. Hou, "Analysis and application of forecasting models in wind power integration: a review of multi-step-ahead wind speed forecasting models," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 960–981, 2016.

[22] E. B. Ssekulima, M. B. Anwar, A. A. Hinai, and M. S. E. Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: a review," *IET Renewable Power Generation*, vol. 10, no. 7, pp. 885–989, 2016.

[23] S. Al-Yahyai, Y. Charabi, and A. Gastli, "Review of the use of numerical weather prediction (NWP) models for wind energy assessment," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 9, pp. 3192–3198, 2010.

[24] Z. Yao, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, no. 5, pp. 255–270, 2014.

[25] J. Shi, J. Guo, and S. Zheng, "Evaluation of hybrid forecasting approaches for wind speed and power generation time series," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3471–3480, 2012.

[26] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, 2015.

[27] J. Wan, J. Liu, G. Ren, and Q. Hu, "Day-ahead prediction of wind speed with deep feature learning," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 30, no. 5, pp. 1–20, 2016.

[28] Q. Hu, R. Zhang, and Y. Zhou, "Transfer learning for short-term wind speed prediction with deep neural networks," *Renewable Energy*, vol. 85, pp. 83–95, 2016.

[29] H. Bludszuweit, J. A. Domínguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 983–991, 2008.

[30] J. H. Li, J. M. Li, J. Y. Wen, and C. Yue, "Generating wind power time series based on its persistence and variation

[31] E. H. Markee, "On the relationships of range to standard deviation of wind fluctuations," *Monthly Weather Review*, vol. 91, no. 2, pp. 83–87, 1963.

[32] G. Ren, J. Liu, and J. Wan, "Prediction of the wind speed turbulence standard deviation," *Journal of Environmental Informatics*, vol. 32, no. 1, pp. 1–13, 2018.

[33] D. C. Hill, D. Mcmillan, K. R. W. Bell, and D. Infield, "Application of auto-regressive models to U.K. Wind speed data for power system impact studies," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 134–141, 2011.

[34] S. Jerez, R. M. Trigo, A. Sarsa, and J. P. Montavez, "Spatio-temporal complementarity between solar and wind power in the iberian Peninsula," *Energy Procedia*, vol. 40, pp. 48–57, 2013.

[35] R. J. Barthelmie, S. T. Frandsen, M. N. Nielsen, and S. C. Pryor, "Modelling and measurements of power losses and turbulence intensity in wind turbine wakes at Middel grunden offshore wind farm," *Wind Energy*, vol. 10, no. 6, pp. 517–528, 2010.

[36] W. G. Früh, A. C. W. Creech, and A. E. Maguire, "Turbulence characteristics in offshore wind farms from LES simulations of lillgrund wind farm," *Energy Procedia*, vol. 59, pp. 182–189, 2014.

[37] G. Gualtieri, "Surface turbulence intensity as a predictor of extrapolated wind resource to the turbine hub height," *Renewable Energy*, vol. 78, pp. 68–81, 2015.

[38] X. Geng, L. Zhou, J. M. Freedman, and M. C. Cervarich, "A case study of effects of atmospheric boundary layer turbulence, wind speed, and stability on wind farm induced temperature changes using observations from a field campaign," *Climate Dynamics*, vol. 46, no. 7-8, pp. 2179–2196, 2016.

[39] L. M. Bardal and L. R. STran, "Influence of turbulence intensity on wind turbine power curves," *Energy Procedia*, vol. 137, pp. 553–558, 2017.

[40] Y. Guo, Q. Wang, D. Zhang, and J. Yu, "A stochastic-process-based method for assessing frequency regulation ability of power systems with wind power fluctuations," *Journal of Environmental Informatics*, vol. 32, no. 1, pp. 45–54, 2018.

[41] J. Liu, G. Ren, J. Wan, Y. Guo, and D. Yu, "Variogram time-series analysis of wind speed," *Renewable Energy*, vol. 99, pp. 483–491, 2016.

[42] E. Cheynet, J. B. Jakobsen, and C. Obhrai, "Spectral characteristics of surface-layer turbulence in the North Sea," *Energy Procedia*, vol. 137, pp. 414–427, 2017.

[43] G. Ren, J. Liu, J. Wan, F. Li, and D. Yu, "The analysis of turbulence intensity based on wind speed data in onshore wind farms," *Renewable Energy*, vol. 123, pp. 756–766, 2018.

[44] A. Cosseron, U. B. Gunturu, and C. A. Schlosser, "Characterization of the wind power resource in europe and its intermittency," *Energy Procedia*, vol. 40, pp. 58–66, 2013.

[45] G. Ren, J. Liu, J. Wan, Y. Guo, and J. Liu, "Measurement and statistical analysis of wind speed intermittency," *Energy*, vol. 118, pp. 632–648, 2016.

[46] G. Ren, J. Wan, J. Liu, and D. Yu, "Analysis of wind power intermittency based on historical wind power data," *Energy*, vol. 150, pp. 482–492, 2018.

[47] M. Qu, J. Wan, and X. Hao, "Analysis of diurnal air temperature range change in the continental United States," *Weather & Climate Extremes*, vol. 4, pp. 86–95, 2014.

[48] H. Chien, H. Y. Cheng, K. H. Yang, and W. T. Chang, "Diurnal and semidiurnal variability of coastal wind over

characteristics," *Science China Technological Sciences*, vol. 57, no. 12, pp. 2475–2486, 2014.

Taiwanese waters," *Wind Energy*, vol. 18, no. 8, pp. 1353–1370, 2015.

[49] S. Wharton and J. K. Lundquist, *Assessing Atmospheric Stability and the Impacts on Wind Characteristics at an Onshore Wind Farm*, Lawrence Livermore National Lab.(LLNL), Livermore, CA, USA, 2010.

[50] C. M. Smith, R. J. Barthelmie, and S. C. Pryor, "In situ observations of the influence of a large onshore wind farm on near-surface temperature, turbulence intensity and wind speed profiles," *Environmental Research Letters*, vol. 8, no. 3, Article ID 034006, 2013.

[51] R. T. Akinnubi and M. O. Adeniyi, "Modeling of diurnal pattern of air temperature in a tropical environment: ile-Ife and Ibadan, Nigeria," *Modeling Earth Systems & Environment*, vol. 3, no. 2, pp. 1–19, 2018.

[52] H. Kanamori, T. Yasunari, and K. Kuraji, "Modulation of the diurnal cycle of rainfall associated with the MJO observed by a dense hourly rain gauge network at sarawak, borneo," *Journal of Climate*, vol. 26, no. 13, pp. 4858–4875, 2013.

[53] C. Li, H. Peng, and J. Sun, "Predictive control and sizing of energy storage to mitigate wind power intermittency," *Wind Energy*, vol. 19, no. 3, pp. 437–451, 2016.

[54] J. Wan, S. Fan, K. Yao, K. Han, E. Peng, and J. Yu, "Spatiotemporal diurnal modulation characteristic of wind speed and power generation revealed by its actual measurement data information processing," *Research Square*, 2021.

Research Article

Analysis on the Differences of Combination Effects of Science and Technology Innovation Policies

Meirong Zhou,[1] **Ping Wei,**[2] **Xi Zeng,**[1] **and Lianbing Deng** ⓘ[1,2]

[1]*Zhuhai Da Hengqin Science and Technology Development Co., Ltd., Zhuhai 519031, China*
[2]*School of Economics, Huazhong University of Science and Technology, Wuhan 430000, China*

Correspondence should be addressed to Lianbing Deng; dhqscience@dhqtech.com

Science and technology innovation (STI) policy is a strategic principle to guide the whole cause of STI. The study on STI policy and its effect is particularly important. Most of the existing studies on the effect of STI policy focus on the effect of a single policy, and the studies on the effect of policy combination and its differences need to be further enriched and improved. This study proposes a method combining system simulation experiment and analysis of variance (ANOVA) to study the differences of combination effects of STI policies. The results show that there are significant effect differences in the combination of STI policies as a whole, but when it comes to different combinations of STI policies, not all policy combinations have significant differences. This study not only points out whether there are significant differences in a certain effect among which combinations of STI policies but also points out whether there are significant differences in all effects among which combinations of STI policies at the same time. This study has theoretical and practical significance for realizing scientific policy-making and sustainable development.

1. Introduction

STI is the inexhaustible driving force for the development of modern countries and the fundamental driving force for realizing sustainable economic development. At present, the improvement of STI ability and the high-quality development of STI have become the theme of the times. Nowadays, countries all over the world attach great importance to their own STI. Effective innovation and high-level patented technology have become important indicators to measure the level of national economic development and comprehensive national strength. In order to accelerate the pace of STI, countries all over the world have formulated STI policies to promote the development of innovation activities and the transformation of achievements. STI policy is the basic action criterion stipulated by a country to realize the STI task in a certain historical period. It is the strategic principle to determine the development direction of STI and guide the whole STI [1]. STI policy can ensure the effective implementation of various innovation activities and the rational allocation of innovation resources. Its reasonable formulation will help to improve innovation performance, so as to promote the construction of national innovation system. As a key factor to strengthen the national scientific and technological strategic force and drive the development of STI, STI policy has attracted extensive attention in the academic circles and has become an important research topic [2].

After studying the existing literature, it is found that the research on policy effects (including STI policy effects) related to this research is mainly reflected in the following two aspects:

(1) What methods are used to study policy effects? The existing research methods on policy effects (including STI policy effects) are divided into two categories: qualitative research methods and quantitative research methods. Most of them analyze the impact of policies from a qualitative perspective. For

example, Surianto et al. studied why policy efforts to reduce disaster risk often fail to improve future disaster relief [3]; Krivosheev et al. studied the effectiveness of EU policy in Walloon, Belgium, and the economic development of the region compared with Flanders, another Belgian region [4]. Based on the current situation of industrial transfer in Guangdong Province, Shi et al. discussed the role of policies in promoting industrial transfer and existing problems, and put forward suggestions on the innovation of regional policies in Guangdong Province [5]; Yu [6] and Lan et al. [7] studied the impact of policies on regional economic development and believed that the government should formulate regional economic policies in line with its own regional economic development direction according to the economic development of each region, so as to scientifically and reasonably allocate the resources of each region; Lin studied the impact of EU policies on Portugal's national STI capacity [8]. In addition, a few of them analyze the impact of policies from a quantitative perspective by establishing models. Wang et al. used system dynamics to build a mechanism model of the impact of STI policies on regional innovation capability and studied the impact mechanism of different policies on regional innovation capability [9]; Avdiushchenko and Zając proposed a set of possible indicators to evaluate the progress of EU countries in realizing circular economy at the regional level [10]; O'Brien and Burrows used vertical and mixed methods to analyze the effectiveness of regional policies for large-scale layoffs [11]; Gülal and Ayaita took the introduction of minimum wage in Germany as the standard experiment and used the double difference model to analyze the impact of minimum wage on well-being [12].

(2) Research on the effect of single policy: most of the existing research on the effect of STI policy is to study the effect of a single STI policy. Cheng et al. found that government research and development (R&D) investment and human investment have a positive effect on STI, while enterprise financing policy has a significant negative effect on STI [13]; Deng and Long [14] and Yu [15] pointed out that R&D investment has a positive impact on STI; Fan [16] and Wang and Deng [17] believed that intellectual property protection helps promote STI; Zhuang et al. emphasized that the environmental policy of STI has a significant impact on STI [18]; Álvarez-Ayuso et al. studied the impact of tax credit on R&D [19]; Raffaello and Paolo studied the innovation effect of R&D subsidy policy in northern Italy [20]; Mukherjee et al. studied the relationship between taxation and innovation and pointed out that taxation hinders innovation [21]; Cappelen et al. believed that tax incentives have little effect on the marketization of new products

and patented products, but they are conducive to the innovation of product production process and the formation of new products [22]; Guerzoni and Raiteri took enterprises in EU Member States as samples, and considered that government procurement not only stimulates enterprises' R&D investment, but also stimulates innovation output [23]; Aschhoff and Sofka empirically tested the effect of government procurement on innovation output by using the survey data of German enterprises and the results show that government procurement can promote enterprise innovation [24]; Song and Zhang studied the relationship between government procurement policies and implementation rules to promote independent innovation, combed the practices in typical countries, and pointed out that government procurement is an important policy tool to support the innovation [25]; Hu et al. used provincial panel data to conduct an empirical test on whether China's government procurement has produced the expected technological innovation effect and the results show that China's government procurement not only does not promote innovation, but hinders innovation [26]. At present, there are also a few studies on the combination effect of STI policies. Kalcheva et al. used the triple difference method to study the impact of the combination of supply policy and demand policy on innovation [27]; Dou et al. found that the combination of STI policies has a significant impact on enterprise technological innovation [28]; Guo et al. believed that the combination effect of various STI policies is better than a single type of policies. When the policies are issued, we should pay attention to the complementarity and systematicness of the policies and avoid focusing too much on a certain link [29].

To sum up, the existing research on the effect of STI policy is mainly the effect of a single policy, and the research on the combination effect of STI policy needs to be further enriched and improved. In fact, the effect of STI policy is often the result of the joint force of multiple policies. At the same time, the existing research on the combination effect of STI policy needs to pay more attention to the research on the difference of policy combination effect. The detailed analysis of the difference of policy combination effect is helpful for researchers and policy makers to have a deeper understanding of the effect of policy combination, and it is also helpful for scientific policy implementation. In view of this, this study will use the ANOVA method to study the differences of the combination effects of STI policies.

The above is the first part of this study, that is, the introduction. This study will be carried out according to the following structure: firstly, the research methods used are introduced, that is, the one-way ANOVA method and multiple comparison method are introduced; secondly, the differences of policy portfolio effects are analyzed in detail by

using one-way ANOVA and multiple comparison method; the last part is the summary of this study.

2. Relevant Theoretical Basis

Because STI has the characteristics of externality and risk of public goods, it will hinder the enthusiasm of the main body of STI, and lead to the fact that the supply quantity in the market is often difficult to reach the social optimal level. This provides a theoretical basis for the government to stimulate the innovation power of STI subjects through policies [30]. Market failure theory holds that in the case of market failure, it is difficult to achieve Pareto optimal state only by market domination. At this time, it is necessary to combine the market with the government. R&D and innovation activities have the characteristics of externality and risk of public goods to a certain extent, which will lead to the failure of STI market, and provide a theoretical basis for the government to intervene in the R&D activities of STI subjects. Endogenous growth theory is based on Schumpeter's innovation theory, takes technological progress as an endogenous variable, and believes that human capital and knowledge accumulation bring the increasing marginal rate of return, so as to ensure economic growth and social progress. These endogenous factors that bring about the economic growth will be affected by government policies and are sensitive to policies. Romer [31] pointed out that we should vigorously promote various policies to promote innovation, such as giving direct subsidies to knowledge output, paying attention to education, strengthening intellectual property protection, etc. Through these policy orientations, we should send signals to STI subjects and stimulate their R&D enthusiasm.

3. Research Method: One-Way ANOVA and Multiple Comparison

ANOVA is used to test the significance of the difference between the mean of two or more samples. Its basic idea is as follows: by analyzing and decomposing the fluctuation of experimental data, and then comparing the possible systematic fluctuation and random fluctuation between each group of experimental data under a certain influencing factor, it is inferred whether there is a significant difference between the overall mean values. If there is a significant difference, it indicates that the influence of this factor is significant; otherwise, it is not significant [32, 33]. When the influence of only one factor on the experimental results is considered, it is called one-way ANOVA; when two factors affect the experimental results, it is called two-way ANOVA; by analogy, when more than two factors affect the experimental results, it is called multiway ANOVA.

The multiple comparison method in the parameter method is to test which overall mean values are equal and which overall mean values are different through pairwise comparison between the overall mean values. There are many methods for multiple comparison. In this study, the least significant difference (LSD) method proposed by Fisher is used. It is a simple deformation of t-test. It makes full use of sample information in the calculation of standard error and uniformly estimates a more robust standard error for the mean of all groups [34].

3.1. Related Concepts and Basic Assumptions of One-Way ANOVA. The one-way ANOVA method involves three terms: factor, level, and observation. Their meanings are as follows:

(1) Factor, also known as condition, refers to the object to be tested in the ANOVA

(2) Level, also known as treatment, refers to different values corresponding to factors

(3) Observation, refers to the experimental data obtained at each level of the factor

The one-way ANOVA method has the following basic assumptions:

(1) The assumption of normality is that each population obeys the normal distribution. For each level of factors, the observed values are simple random samples from the normal population.

(2) The homogeneity of variance is assumed, that is, the variance of each population should be equal. For each group of observation data, they are extracted from the normal distribution of the same variance.

(3) Independence assumption, that is, the observations are independent of each other.

3.2. Basic Process of One-Way ANOVA. When the one-way ANOVA method is used for research analysis, the basic process shown in Figure 1 can be adopted.

Firstly, the null hypothesis and alternative hypothesis are put forward. If the factor to be tested has k levels, and the corresponding mean value of each level is u_i, where $i = 1, 2, \ldots, k$, then the null hypothesis and alternative hypothesis are as follows:

H_0: $u_1 = u_2 = \ldots = u_k$ (independent variables have no significant impact on dependent variables);

H_1: $u_1, u_2, \ldots, u_k$ are not all equal (independent variables have a significant impact on dependent variables).

The null hypothesis (H_0) means that the factor to be tested has no impact on the experimental results, while the alternative hypothesis (H_1) means that the factor to be tested has an impact on the experimental results.

Secondly, the test statistics are constructed and calculated. In order to test whether the null hypothesis (H_0) is true, it is necessary to construct appropriate test statistics first. Specifically, it is necessary to construct three sums of squares of error, which are sum of squares of total error (SST), sum of squares of factor error (SSA) and sum of
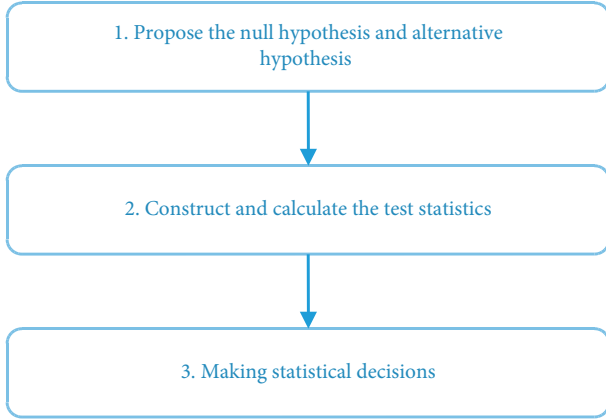
FIGURE 1: Basic process of one-way ANOVA.

squares of random error (SSE). Let x_{ij} represent the j-th observation value of the i-th level, then the calculation formulas of these three sums of squares of error are

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{\bar{x}} \right)^2,$$

$$SSA = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(\bar{x}_i - \bar{\bar{x}} \right)^2 = \sum_{i=1}^{k} n_i \left(\bar{x}_i - \bar{\bar{x}} \right)^2,$$

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i \right)^2, \tag{1}$$

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i},$$

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}}{n} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{n},$$

where $\bar{x}_i$ represents the mean value of the factor at Level i and n_i is the number of the i-th overall experimental data, $i = 1, 2, \ldots, k$; $\bar{x}$ represents the mean value of all observations, $n = n_1 + n_2 + \cdots + n_k$.

These three sums of squares of error satisfy the following identity relationship.

$$SST = SSA + SSE. \tag{2}$$

The sum of squares of error divided by their corresponding degrees of freedom is called mean square error. The degree of freedom of SST is $n - 1$, where n is the number of all observed values. The degree of freedom of SSA is $k - 1$, and its corresponding mean square error is usually recorded as MSA, where k is the number of factor levels. The degree of freedom of SSE is $n - k$, and its corresponding mean square error is usually recorded as MSE. Thus, the value of $F = MSA/MSE$ can be obtained, and it obeys the F $(k - 1, n - k)$ distribution.

Finally, make statistical decisions. According to the given significance level α, check the F distribution table to determine F_α $(k - 1, n - k)$. If $F > F_\alpha$ (or $P < \alpha$), then reject the

null hypothesis, indicating that the factor has a significant impact on the experimental results. Otherwise, accept the null hypothesis and consider that the factor has no significant impact on the experimental results.

3.3. Basic Flow of LSD Method. When the one-way ANOVA method is used for research analysis, the basic process shown in Figure 1 can be adopted.

After one-way ANOVA, if you want to know which population means are equal and which population means are different, you need to make multiple comparisons. In this study, LSD method is used for multiple comparison, and its basic steps are similar to one-way ANOVA, as follows:

Firstly, the null hypothesis and alternative hypothesis are put forward.

H_0: $u_i = u_j$;

H_1: u_i, u_j are not equal where $1 \le i < j \le k$.

Secondly, the test statistics are constructed and calculated. The test statistic of LSD method is t statistic, and its calculation formula is

$$t = \frac{\left(\bar{x}_i - \bar{x}_j \right) - \left(u_i - u_j \right)}{\sqrt{MSE \left(1/n_i + 1/n_j \right)}}. \tag{3}$$

t statistic obeys the t distribution with $n - k$ degrees of freedom.

Finally, make statistical decisions. According to the given significance level α, check the t distribution table to determine $t_{\alpha/2}$ $(n - k)$. If $|t| > t_{\alpha/2}(n - k)$ (or $P < \alpha$), then reject the null hypothesis, indicating that there is a difference between the two population mean values. Otherwise, accept the null hypothesis and consider that there is no difference between the two population mean values.

3.4. Welch's ANOVA and Games–Howell Test. It is worth noting that in this study, when the variances are unequal, we will use Welch's ANOVA method to replace the aforementioned one-way ANOVA method for corresponding analysis, because Welch statistic is better than F statistic when the variance is unequal [35]. At the same time, we will use the Games–Howell test method to replace the above LSD method for corresponding analysis. For a detailed introduction to the Games–Howell test method, please refer to reference [36].

4. Analysis on the Differences of Combination Effects

This study is based on our previous research achievement "Analysis on the combination effect of science and technology innovation policy: from the perspective of system simulation" (hereinafter referred to as Achievement A), which proposes a system simulation method to analyze the combination effect of STI policy. This study uses the system simulation model in Achievement A to carry out simulation experiments and obtain simulation experimental data, so as to analyze the differences of policy combination effects. The

following is a brief introduction to the system simulation model in Achievement A.

Rothwell and Zegveld [37] believe that the policy effect originates from different levels. Therefore, the STI policy can be divided into three categories, one is supply policy, and the other two are demand policy and environmental policy. In this study, our goal is to analyze the impact of different level STI policies. Therefore, this classification method will be adopted. Considering the quantification of policies and the availability of data, the STI supply policy in Achievement A only considers the capital investment policy and talent training policy, the STI demand policy only considers the intellectual property policy and STI service policy, and the STI environment policy only considers the enterprise financing policy and enterprise tax policy. The Achievement A establishes the conceptual model as shown in Figure 2, that is, the causality diagram. The system simulation model corresponding to the conceptual model, that is, the stock flow diagram, is shown in Figure 3.

The above is a brief introduction to the system simulation model in Achievement A. For the detailed introduction of Achievement A, please refer to my previous research Achievement A.

It is worth noting that Achievement A has conducted an empirical study using the relevant data of STI in Guangdong Province from 2010 to 2019. Its system simulation model does not consider the uncertain factors, but this study will take the uncertain factors into account. Specifically, this study will add random distribution to the expression of relevant variables in Achievement A, as shown in Table 1. For example, for the annual GDP, the expression in this study is: annual GDP = the expression of annual GDP in Achievement A + RANDOM NORMAL (−5460.12, 5291.63, 0, 3622.37). Among them, RANDOM NORMAL(min, max, mean, stdev) is the normal distribution form in the system dynamics simulation software VENSIM, which needs to input four parameters.

It can be seen from Table 1 that all the random distributions in the expressions of Variable 1 to 9 are a normal distribution, because the expressions corresponding to these variables in Achievement A are obtained by linear regression analysis method, which requires that the error term obey a normal distribution and the mean value is 0. In Achievement A, the expressions of Variable 10 to 13 are obtained by weighted regression analysis. This study assumes that the error terms of variables 10 to 13 obey a normal distribution. The parameters of the random distribution shown in Table 1 are obtained by analyzing the error term.

This study will make a difference analysis on the combination effects of the above three types of policies in SPSS statistical analysis software. The effects here mainly refer to the economic effect and STI effect. The per capita GDP and patent application acceptance are used as the effect indicators respectively. In this study, each type of policy takes three levels. Specifically, the three levels of supply policy are benchmark level, relative benchmark increase of 5% and relative benchmark decrease of 5%. The three levels of demand policy are similar to those of supply policy. The three levels of environmental policy are the benchmark level, the

relative benchmark increase of 1% and the relative benchmark decrease of 1%. The benchmark levels for each policy are detailed in Achievement A. Since there are three types of policies, and each type of policy has three levels, this study will design $3^3 = 27$ simulation experiments, as shown in Table 2.

4.1. Descriptive Statistics.
After the preliminary analysis of the data (that is, the model output data with the simulation time of 2025) obtained from the above simulation experiments, some descriptive statistical analysis results are obtained as shown in Table 3. It is not difficult to find from Table 3 that the number of experiments in each experiment is the same, 50 times. From Table 3, we can also see the mean value and standard deviation of per capita GDP (PCGDP) and annual patent application acceptance (APAA) of each simulation experiment.

4.2. Test of Normality and Test of Homogeneity of Variance.
Because the one-way ANOVA method requires that the population should meet the normality and homogeneity of variance, it is necessary to test the normality and homogeneity of variance before analyzing the differences of policy combination effects. It can be seen from Table 1 that the normality requirements are met. From Table 4, it can be seen that the significance (P value) of per capita GDP is greater than 0.05. Thus, the homogeneity of variance is met at this time. For the annual patent application acceptance, the significance (P value) is less than 0.05. Hence, the homogeneity of variance is not met at this time.

4.3. Difference Analysis of Economic Effect.
For the per capita GDP, since each population meets the requirements of normality and homogeneity of variance, we use the abovementioned one-way ANOVA method for analysis. As can be seen from Table 5, the significance (P value) is less than 0.05, which indicates that there are significant differences in the economic effect (i.e., per capita GDP) of various combinations of STI policies. It is worth noting that this significant difference refers to the significant difference of the whole. At this time, we do not know which combinations have significant differences, and which combinations have no significant differences. In order to answer this question, it is necessary to make multiple comparisons so that we can clearly understand which combinations have significant differences. It can be seen from Table 6 that when 0.05 is taken as the significance level, there is a significant difference between the economic effect of the policy combination corresponding to Experiment 1 and the economic effect of other policy combinations, except that it is not significantly different from the economic effect of the policy combination corresponding to Experiment 7, 10, 13, 19, and 25. In addition to the policy combination corresponding to Experiment 1, similar difference analysis can also be carried out for other policy combinations in combination with Table 6, which will not be repeated here.
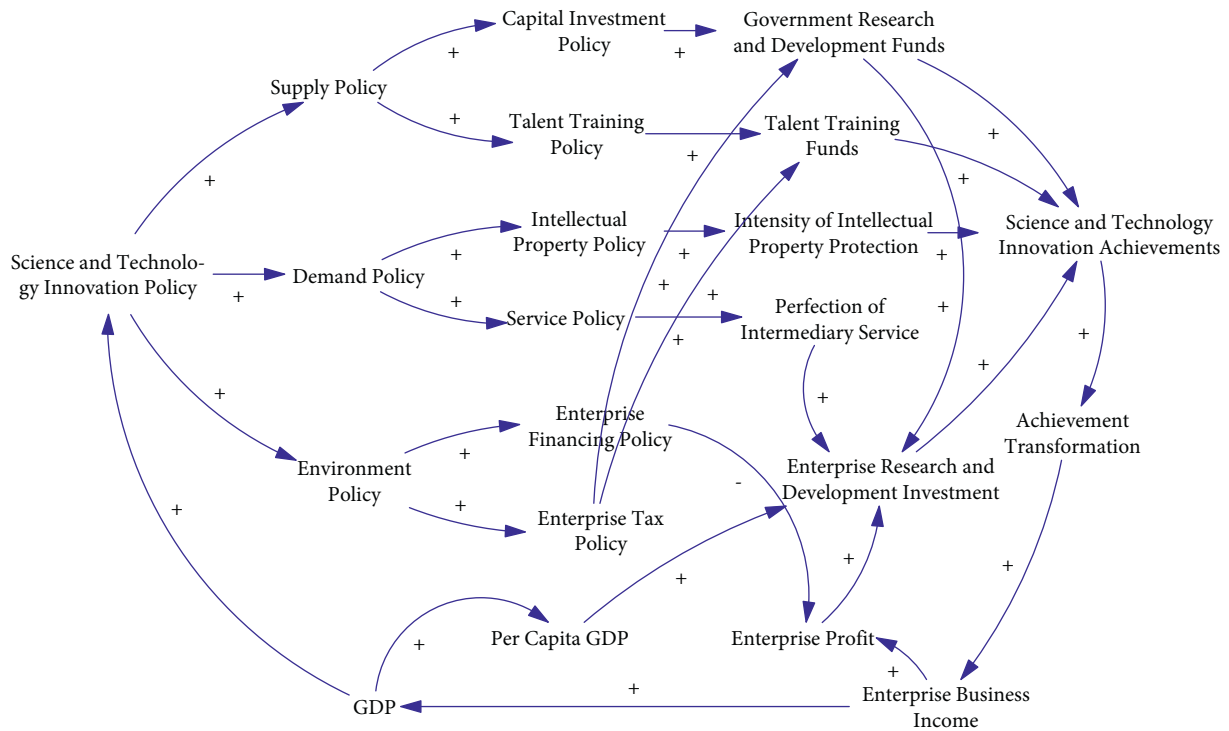
FIGURE 2: Conceptual model.
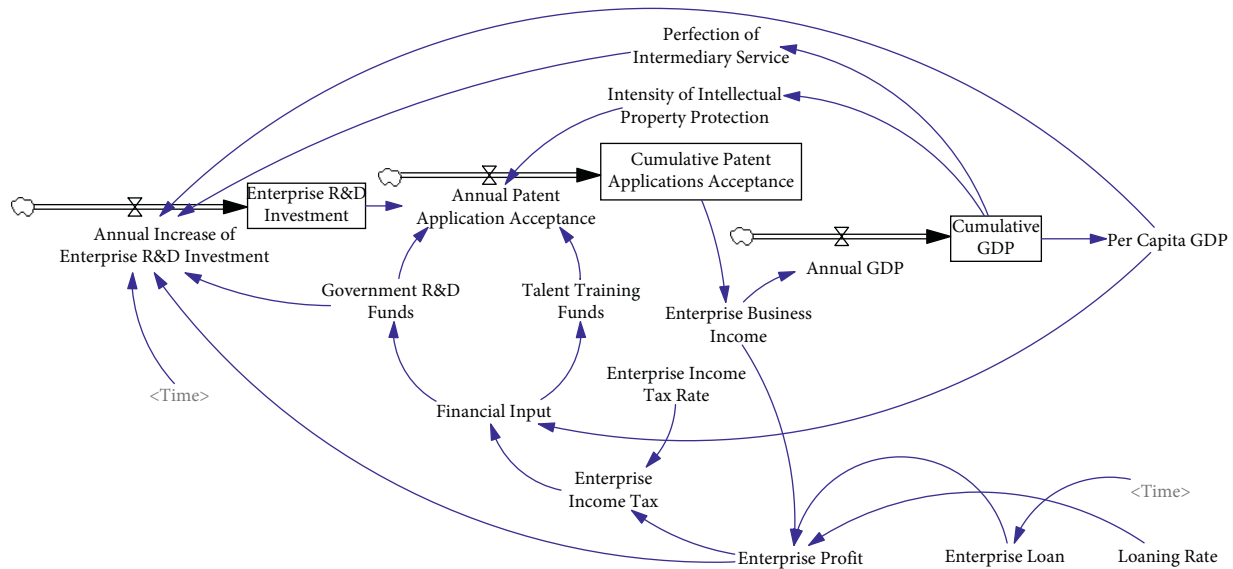
FIGURE 3: Simulation model.

4.4. Difference Analysis of STI Effect. For the annual patent application acceptance, although each population meets the requirements of normality, they do not meet the requirements of homogeneity of variance. Therefore, the abovementioned one-way ANOVA method is not applicable here. At this time, we can use the abovementioned Welch's ANOVA method for analysis. As can be seen from Table 7, the significance (*P* value) is less than 0.05, which shows that the STI effect (i.e., annual patent application acceptance) of various combinations of STI policies are significantly

different as a whole. In order to further understand which population mean values are equal and which population mean values are different, it is necessary to make multiple comparisons using the aforementioned Games–Howell test method. It is not difficult to find from Table 8 that, at the significance level of 0.05, there is no significant difference between the STI effect of the policy combination corresponding to Experiment 1 and the STI effect of the policy combination corresponding to Experiment 4, 7, 10, 13, 16, 19, 22, and 25. However, there are significant differences

TABLE 1: Random distribution in variable expressions.

Number	Variable name	Random distribution to be added
Variable 1	Annual GDP	RANDOM NORMAL (−5460.12, 5291.63, 0, 3622.37)
Variable 2	Enterprise loan	RANDOM NORMAL (−6092.92, 8022.4, 0, 5002.37)
Variable 3	Enterprise income tax	RANDOM NORMAL (−107.338, 67.0867, 0, 52.0629)
Variable 4	Perfection of intermediary service	RANDOM NORMAL (−1.2481, 0.755273, 0, 0.789901)
Variable 5	Enterprise business income	RANDOM NORMAL (−12078.2, 13275.6, 0, 8515.95)
Variable 6	Talent training funds	RANDOM NORMAL (−213.521, 161.98, 0, 126.222)
Variable 7	Government R&D funds	RANDOM NORMAL (−118.463, 103.843, 0, 76.0717)
Variable 8	Intensity of intellectual property protection	RANDOM NORMAL (−0.184447, 0.11458, 0, 0.117456)
Variable 9	Per capita GDP	RANDOM NORMAL (−1081.07, 1585.02, 0, 838.532)
Variable 10	Annual increase of enterprise R&D investment	RANDOM NORMAL (−71.0683, 19.1128, −8.75484, 25.325)
Variable 11	Annual patent application acceptance	RANDOM NORMAL (−129351, 50584.4, −6503, 46332)
Variable 12	Enterprise profit	RANDOM NORMAL (−362.065, 1172.9, 77.7868, 427.839)
Variable 13	Financial input	RANDOM NORMAL (−661.539, 1540.54, 66.5193, 658.025)

TABLE 2: Simulation experiment.

Experiment	Supply policy	Demand policy	Environmental policy
Experiment 1	Benchmark	Benchmark	Benchmark
Experiment 2	Increase of 5%	Benchmark	Benchmark
Experiment 3	Decrease of 5%	Benchmark	Benchmark
Experiment 4	Benchmark	Increase of 5%	Benchmark
Experiment 5	Increase of 5%	Increase of 5%	Benchmark
Experiment 6	Decrease of 5%	Increase of 5%	Benchmark
Experiment 7	Benchmark	Decrease of 5%	Benchmark
Experiment 8	Increase of 5%	Decrease of 5%	Benchmark
Experiment 9	Decrease of 5%	Decrease of 5%	Benchmark
Experiment 10	Benchmark	Benchmark	Increase of 1%
Experiment 11	Increase of 5%	Benchmark	Increase of 1%
Experiment 12	Decrease of 5%	Benchmark	Increase of 1%
Experiment 13	Benchmark	Increase of 5%	Increase of 1%
Experiment 14	Increase of 5%	Increase of 5%	Increase of 1%
Experiment 15	Decrease of 5%	Increase of 5%	Increase of 1%
Experiment 16	Benchmark	Decrease of 5%	Increase of 1%
Experiment 17	Increase of 5%	Decrease of 5%	Increase of 1%
Experiment 18	Decrease of 5%	Decrease of 5%	Increase of 1%
Experiment 19	Benchmark	Benchmark	Decrease of 1%
Experiment 20	Increase of 5%	Benchmark	Decrease of 1%
Experiment 21	Decrease of 5%	Benchmark	Decrease of 1%
Experiment 22	Benchmark	Increase of 5%	Decrease of 1%
Experiment 23	Increase of 5%	Increase of 5%	Decrease of 1%
Experiment 24	Decrease of 5%	Increase of 5%	Decrease of 1%
Experiment 25	Benchmark	Decrease of 5%	Decrease of 1%
Experiment 26	Increase of 5%	Decrease of 5%	Decrease of 1%
Experiment 27	Decrease of 5%	Decrease of 5%	Decrease of 1%

with other policy combinations in the effect of STI. In addition to the policy combination corresponding to Experiment 1, similar difference analysis can also be carried out for other policy combinations in combination with Table 8, which will not be repeated here.

4.5. Comprehensive Analysis of Differences. The above analyzes the differences of the economic effect and STI effect of each combination of STI policies. Next, we take policy Combination 1 as an example and analyze these two effects of each combination of STI policies at the same time. By sorting out Tables 6 and 8, we can get Table 9. It can be seen from Table 9 that the economic effect and STI effect of policy Combination 1 are significantly different from those of policy Combination 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 21, 23, 24, 26 and 27, but not from those of policy Combination 7, 10, 13, 19 and 25. In addition, there are significant differences in economic effect between policy Combination 1 and policy Combination 4, 16 and 22, but there is no significant difference in STI effect.

TABLE 3: Descriptive statistical analysis of effect indicators.

		N	Mean	Std. dev	95% CI for mean		Min	Max
					LB	UB		
PCGDP	1	50	116875.19	2756.01	116091.94	117658.44	110202.14	122390.17
	2	50	119737.81	3000.28	118885.14	120590.48	112699.75	126021.47
	3	50	114350.49	2552.16	113625.18	115075.81	107994.00	119207.41
	4	50	117970.75	2781.64	117180.22	118761.28	111239.44	123558.27
	5	50	120891.43	3030.46	120030.19	121752.68	113788.03	127260.36
	6	50	115395.78	2574.09	114664.23	116127.32	108986.92	120314.95
	7	50	115786.38	2730.63	115010.35	116562.42	109171.17	121229.48
	8	50	118591.91	2970.42	117747.72	119436.09	111618.66	124791.30
	9	50	113311.15	2530.45	112592.00	114030.30	107006.61	118106.33
	10	50	116676.04	2769.68	115888.91	117463.17	109978.14	122227.02
	11	50	119551.25	3019.44	118693.13	120409.36	112478.59	125886.61
	12	50	114144.73	2561.82	113416.67	114872.79	107770.75	119026.17
	13	50	117778.65	2796.14	116983.99	118573.30	111020.88	123404.47
	14	50	120713.78	3050.69	119846.79	121580.78	113573.70	127137.38
	15	50	115195.57	2584.37	114461.10	115930.04	108768.08	120140.98
	16	50	115580.43	2743.49	114800.74	116360.12	108941.84	121057.23
	17	50	118396.74	2988.53	117547.41	119246.07	111390.78	124644.86
	18	50	113100.02	2539.49	112378.31	113821.73	106779.22	117917.95
	19	50	117037.61	2740.57	116258.75	117816.47	110393.23	122511.11
	20	50	119880.62	2978.97	119034.00	120727.23	112881.50	126106.11
	21	50	114525.60	2541.09	113803.43	115247.77	108189.75	119353.38
	22	50	118125.65	2765.38	117339.73	118911.56	111424.56	123669.33
	23	50	121024.80	3008.07	120169.91	121879.68	113962.31	127332.38
	24	50	115564.91	2562.37	114836.69	116293.13	109177.98	120453.05
	25	50	115956.06	2715.99	115184.18	116727.93	109367.95	121359.86
	26	50	118743.84	2950.15	117905.41	119582.26	111807.53	124887.91
	27	50	113492.03	2519.99	112775.86	114208.20	107206.84	118259.78
	Total	1350	116977.75	3635.29	116783.65	117171.84	106779.22	127332.38
APAA	1	50	5447216.71	566848.37	5286120.19	5608313.23	4123422.50	6897111.50
	2	50	6781777.86	726435.97	6575327.04	6988228.68	5128423.50	8586776.00
	3	50	4411721.91	449265.41	4284042.09	4539401.73	3340311.75	5586297.50
	4	50	5727187.33	585571.63	5560769.71	5893604.95	4358983.50	7220039.00
	5	50	7124829.99	751794.28	6911172.42	7338487.56	5413846.50	8985734.00
	6	50	4643638.21	463370.31	4511949.82	4775326.59	3537511.25	5851589.00
	7	50	5174886.30	548610.50	5018972.91	5330799.68	3894416.50	6582780.00
	8	50	6448667.84	701800.64	6249218.30	6648117.38	4851419.00	8199133.00
	9	50	4185764.68	435493.88	4061998.69	4309530.67	3148286.50	5327639.00
	10	50	5452400.10	577314.29	5288329.19	5616471.01	4106576.00	6928417.00
	11	50	6814753.65	743851.16	6603353.49	7026153.81	5126440.50	8659248.00
	12	50	4400640.36	455516.10	4271184.11	4530096.60	3315451.00	5592566.00
	13	50	5739246.96	597028.50	5569573.34	5908920.58	4346903.50	7260298.50
	14	50	7167935.57	770751.97	6948890.28	7386980.86	5418855.50	9071408.00
	15	50	4637260.98	470270.14	4503611.68	4770910.28	3515921.25	5864043.00
	16	50	5173637.43	558134.43	5015017.38	5332257.49	3873134.75	6605654.50
	17	50	6472158.41	717754.95	6268174.71	6676142.11	4842897.00	8259190.00
	18	50	4170287.47	441123.80	4044921.47	4295653.47	3120412.50	5328120.50
	19	50	5428011.83	555576.99	5270118.60	5585905.06	4128479.50	6850074.50
	20	50	6730683.09	708049.78	6529457.57	6931908.61	5115007.50	8494183.00
	21	50	4412000.78	442373.58	4286279.60	4537721.96	3356138.50	5567822.00
	22	50	5700801.13	573329.26	5537862.76	5863739.50	4358943.00	7163762.50
	23	50	7063271.32	731911.68	6855264.32	7271278.32	5393022.50	8879698.00
	24	50	4638950.83	455836.91	4509403.41	4768498.24	3549818.75	5826734.00
	25	50	5162424.61	538262.52	5009452.09	5315397.13	3904222.50	6544449.00
	26	50	6407408.94	684837.86	6212780.17	6602037.71	4844911.50	8119192.50
	27	50	4190705.05	429214.88	4068723.53	4312686.57	3167381.50	5315209.50
	Total	1350	5544750.72	1157255.99	5482963.24	5606538.19	3120412.50	9071408.00

TABLE 4: Test of homogeneity of variances.

		Levene statistic	df1	df2	Sig.
PCGDP	Based on mean	0.479	26	1323	0.988
	Based on median	0.467	26	1323	0.990
	Based on median and with adjusted df	0.467	26	1302.907	0.990
	Based on trimmed mean	0.475	26	1323	0.989
APAA	Based on mean	3.182	26	1323	0.000
	Based on median	3.129	26	1323	0.000
	Based on median and with adjusted df	3.129	26	1116.829	0.000
	Based on trimmed mean	3.189	26	1323	0.000

TABLE 5: ANOVA.

			Per_Capita_GDP		
	Sum of squares	df	Mean square	F	Sig.
Between groups	7638617410.110	26	293792977.312	38.148	0.000
Within groups	10188907629.601	1323	7701366.311		
Total	17827525039.710	1349			

TABLE 6: Multiple comparisons on dependent variable Per_Capita_GDP.

(I) experiment	(J) experiment	Mean difference ($I-J$)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
1	2	−2862.62*	555.03	0.000	−3951.45	−1773.79
	3	2524.70*	555.03	0.000	1435.87	3613.53
	4	−1095.56*	555.03	0.049	−2184.39	−6.73
	5	−4016.24*	555.03	0.000	−5105.07	−2927.41
	6	1479.42*	555.03	0.008	390.59	2568.25
	7	1088.81	555.03	0.050	−0.02	2177.64
	8	−1716.71*	555.03	0.002	−2805.54	−627.89
	9	3564.04*	555.03	0.000	2475.22	4652.87
	10	199.15	555.03	0.720	−889.68	1287.98
	11	−2676.05*	555.03	0.000	−3764.88	−1587.22
	12	2730.46*	555.03	0.000	1641.63	3819.29
	13	−903.45	555.03	0.104	−1992.28	185.37
	14	−3838.59*	555.03	0.000	−4927.42	−2749.76
	15	1679.62*	555.03	0.003	590.79	2768.45
	16	1294.76*	555.03	0.020	205.93	2383.59
	17	−1521.55*	555.03	0.006	−2610.37	−432.72
	18	3775.17*	555.03	0.000	2686.34	4864.00
	19	−162.42	555.03	0.770	−1251.25	926.41
	20	−3005.42*	555.03	0.000	−4094.25	−1916.59
	21	2349.59*	555.03	0.000	1260.76	3438.42
	22	−1250.45*	555.03	0.024	−2339.28	−161.63
	23	−4149.60*	555.03	0.000	−5238.43	−3060.77
	24	1310.29*	555.03	0.018	221.46	2399.11
	25	919.13	555.03	0.098	−169.69	2007.96
	26	−1868.64*	555.03	0.001	−2957.47	−779.81
	27	3383.16*	555.03	0.000	2294.33	4471.99

TABLE 6: Continued.

(*I*) experiment	(*J*) experiment	Mean difference ($I - J$)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
2	1	2862.62*	555.03	0.000	1773.79	3951.45
	3	5387.32*	555.03	0.000	4298.49	6476.15
	4	1767.06*	555.03	0.001	678.23	2855.89
	5	−1153.62*	555.03	0.038	−2242.45	−64.79
	6	4342.03*	555.03	0.000	3253.21	5430.86
	7	3951.42*	555.03	0.000	2862.60	5040.25
	8	1145.90*	555.03	0.039	57.07	2234.73
	9	6426.66*	555.03	0.000	5337.83	7515.49
	10	3061.77*	555.03	0.000	1972.94	4150.60
	11	186.56	555.03	0.737	−902.26	1275.39
	12	5593.08*	555.03	0.000	4504.25	6681.91
	13	1959.16*	555.03	0.000	870.33	3047.99
	14	−975.97	555.03	0.079	−2064.80	112.86
	15	4542.24*	555.03	0.000	3453.41	5631.07
	16	4157.38*	555.03	0.000	3068.55	5246.20
	17	1341.07*	555.03	0.016	252.24	2429.90
	18	6637.79*	555.03	0.000	5548.96	7726.62
	19	2700.20*	555.03	0.000	1611.37	3789.03
	20	−142.81	555.03	0.797	−1231.63	946.02
	21	5212.21*	555.03	0.000	4123.38	6301.04
	22	1612.16*	555.03	0.004	523.33	2700.99
	23	−1286.99*	555.03	0.021	−2375.81	−198.16
	24	4172.90*	555.03	0.000	3084.07	5261.73
	25	3781.75*	555.03	0.000	2692.92	4870.58
	26	993.97	555.03	0.074	−94.86	2082.80
	27	6245.78*	555.03	0.000	5156.95	7334.61
3	1	−2524.70*	555.03	0.000	−3613.53	−1435.87
	2	−5387.32*	555.03	0.000	−6476.15	−4298.49
	4	−3620.26*	555.03	0.000	−4709.09	−2531.43
	5	−6540.94*	555.03	0.000	−7629.77	−5452.11
	6	−1045.28	555.03	0.060	−2134.11	43.55
	7	−1435.89*	555.03	0.010	−2524.72	−347.06
	8	−4241.41*	555.03	0.000	−5330.24	−3152.59
	9	1039.34	555.03	0.061	−49.48	2128.17
	10	−2325.55*	555.03	0.000	−3414.38	−1236.72
	11	−5200.75*	555.03	0.000	−6289.58	−4111.92
	12	205.76	555.03	0.711	−883.07	1294.59
	13	−3428.15*	555.03	0.000	−4516.98	−2339.33
	14	−6363.29*	555.03	0.000	−7452.12	−5274.46
	15	−845.08	555.03	0.128	−1933.91	243.75
	16	−1229.94*	555.03	0.027	−2318.77	−141.11
	17	−4046.25*	555.03	0.000	−5135.07	−2957.42
	18	1250.47*	555.03	0.024	161.64	2339.30
	19	−2687.12*	555.03	0.000	−3775.95	−1598.29
	20	−5530.12*	555.03	0.000	−6618.95	−4441.29
	21	−175.11	555.03	0.752	−1263.94	913.72
	22	−3775.15*	555.03	0.000	−4863.98	−2686.33
	23	−6674.30*	555.03	0.000	−7763.13	−5585.47
	24	−1214.41*	555.03	0.029	−2303.24	−125.59
	25	−1605.57*	555.03	0.004	−2694.39	−516.74
	26	−4393.34*	555.03	0.000	−5482.17	−3304.51
	27	858.46	555.03	0.122	−230.37	1947.29

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
4	1	1095.56*	555.03	0.049	6.73	2184.39
	2	−1767.06*	555.03	0.001	−2855.89	−678.23
	3	3620.26*	555.03	0.000	2531.43	4709.09
	5	−2920.68*	555.03	0.000	−4009.51	−1831.85
	6	2574.97*	555.03	0.000	1486.15	3663.80
	7	2184.37*	555.03	0.000	1095.54	3273.19
	8	−621.16	555.03	0.263	−1709.98	467.67
	9	4659.60*	555.03	0.000	3570.77	5748.43
	10	1294.71*	555.03	0.020	205.88	2383.54
	11	−1580.50*	555.03	0.004	−2669.32	−491.67
	12	3826.02*	555.03	0.000	2737.19	4914.85
	13	192.10	555.03	0.729	−896.72	1280.93
	14	−2743.03*	555.03	0.000	−3831.86	−1654.20
	15	2775.18*	555.03	0.000	1686.35	3864.01
	16	2390.32*	555.03	0.000	1301.49	3479.15
	17	−425.99	555.03	0.443	−1514.82	662.84
	18	4870.73*	555.03	0.000	3781.90	5959.56
	19	933.14	555.03	0.093	−155.69	2021.97
	20	−1909.87*	555.03	0.001	−2998.69	−821.04
	21	3445.15*	555.03	0.000	2356.32	4533.98
	22	−154.90	555.03	0.780	−1243.72	933.93
	23	−3054.05*	555.03	0.000	−4142.87	−1965.22
	24	2405.84*	555.03	0.000	1317.01	3494.67
	25	2014.69*	555.03	0.000	925.86	3103.52
	26	−773.09	555.03	0.164	−1861.91	315.74
	27	4478.72*	555.03	0.000	3389.89	5567.55
5	1	4016.24*	555.03	0.000	2927.41	5105.07
	2	1153.62*	555.03	0.038	64.79	2242.45
	3	6540.94*	555.03	0.000	5452.11	7629.77
	4	2920.68*	555.03	0.000	1831.85	4009.51
	6	5495.66*	555.03	0.000	4406.83	6584.48
	7	5105.05*	555.03	0.000	4016.22	6193.88
	8	2299.52*	555.03	0.000	1210.70	3388.35
	9	7580.28*	555.03	0.000	6491.45	8669.11
	10	4215.39*	555.03	0.000	3126.56	5304.22
	11	1340.19*	555.03	0.016	251.36	2429.01
	12	6746.70*	555.03	0.000	5657.87	7835.53
	13	3112.78*	555.03	0.000	2023.96	4201.61
	14	177.65	555.03	0.749	−911.18	1266.48
	15	5695.86*	555.03	0.000	4607.03	6784.69
	16	5311.00*	555.03	0.000	4222.17	6399.83
	17	2494.69*	555.03	0.000	1405.86	3583.52
	18	7791.41*	555.03	0.000	6702.58	8880.24
	19	3853.82*	555.03	0.000	2764.99	4942.65
	20	1010.82	555.03	0.069	−78.01	2099.64
	21	6365.83*	555.03	0.000	5277.00	7454.66
	22	2765.78*	555.03	0.000	1676.96	3854.61
	23	−133.36	555.03	0.810	−1222.19	955.46
	24	5326.52*	555.03	0.000	4237.70	6415.35
	25	4935.37*	555.03	0.000	3846.54	6024.20
	26	2147.60*	555.03	0.000	1058.77	3236.42
	27	7399.40*	555.03	0.000	6310.57	8488.23

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
	1	−1479.42*	555.03	0.008	−2568.25	−390.59
	2	−4342.03*	555.03	0.000	−5430.86	−3253.21
	3	1045.28	555.03	0.060	−43.55	2134.11
	4	−2574.97*	555.03	0.000	−3663.80	−1486.15
	5	−5495.66*	555.03	0.000	−6584.48	−4406.83
	7	−390.61	555.03	0.482	−1479.44	698.22
	8	−3196.13*	555.03	0.000	−4284.96	−2107.30
	9	2084.63*	555.03	0.000	995.80	3173.46
	10	−1280.27*	555.03	0.021	−2369.09	−191.44
	11	−4155.47*	555.03	0.000	−5244.30	−3066.64
	12	1251.04*	555.03	0.024	162.21	2339.87
	13	−2382.87*	555.03	0.000	−3471.70	−1294.04
6	14	−5318.01*	555.03	0.000	−6406.83	−4229.18
	15	200.20	555.03	0.718	−888.63	1289.03
	16	−184.66	555.03	0.739	−1273.49	904.17
	17	−3000.96*	555.03	0.000	−4089.79	−1912.13
	18	2295.76*	555.03	0.000	1206.93	3384.58
	19	−1641.84*	555.03	0.003	−2730.66	−553.01
	20	−4484.84*	555.03	0.000	−5573.67	−3396.01
	21	870.18	555.03	0.117	−218.65	1959.00
	22	−2729.87*	555.03	0.000	−3818.70	−1641.04
	23	−5629.02*	555.03	0.000	−6717.85	−4540.19
	24	−169.13	555.03	0.761	−1257.96	919.70
	25	−560.28	555.03	0.313	−1649.11	528.55
	26	−3348.06*	555.03	0.000	−4436.89	−2259.23
	27	1903.75*	555.03	0.001	814.92	2992.57
	1	−1088.81	555.03	0.050	−2177.64	0.02
	2	−3951.42*	555.03	0.000	−5040.25	−2862.60
	3	1435.89*	555.03	0.010	347.06	2524.72
	4	−2184.37*	555.03	0.000	−3273.19	−1095.54
	5	−5105.05*	555.03	0.000	−6193.88	−4016.22
	6	390.61	555.03	0.482	−698.22	1479.44
	8	−2805.52*	555.03	0.000	−3894.35	−1716.69
	9	2475.24*	555.03	0.000	1386.41	3564.06
	10	−889.66	555.03	0.109	−1978.49	199.17
	11	−3764.86*	555.03	0.000	−4853.69	−2676.03
	12	1641.65*	555.03	0.003	552.82	2730.48
	13	−1992.26*	555.03	0.000	−3081.09	−903.43
7	14	−4927.40*	555.03	0.000	−6016.23	−3838.57
	15	590.81	555.03	0.287	−498.02	1679.64
	16	205.95	555.03	0.711	−882.88	1294.78
	17	−2610.35*	555.03	0.000	−3699.18	−1521.52
	18	2686.36*	555.03	0.000	1597.54	3775.19
	19	−1251.23*	555.03	0.024	−2340.06	−162.40
	20	−4094.23*	555.03	0.000	−5183.06	−3005.40
	21	1260.78*	555.03	0.023	171.96	2349.61
	22	−2339.26*	555.03	0.000	−3428.09	−1250.43
	23	−5238.41*	555.03	0.000	−6327.24	−4149.58
	24	221.48	555.03	0.690	−867.35	1310.31
	25	−169.67	555.03	0.760	−1258.50	919.16
	26	−2957.45*	555.03	0.000	−4046.28	−1868.62
	27	2294.35*	555.03	0.000	1205.53	3383.18

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
8	1	1716.71*	555.03	0.002	627.89	2805.54
	2	−1145.90*	555.03	0.039	−2234.73	−57.07
	3	4241.41*	555.03	0.000	3152.59	5330.24
	4	621.16	555.03	0.263	−467.67	1709.98
	5	−2299.52*	555.03	0.000	−3388.35	−1210.70
	6	3196.13*	555.03	0.000	2107.30	4284.96
	7	2805.52*	555.03	0.000	1716.69	3894.35
	9	5280.76*	555.03	0.000	4191.93	6369.59
	10	1915.87*	555.03	0.001	827.04	3004.69
	11	−959.34	555.03	0.084	−2048.17	129.49
	12	4447.17*	555.03	0.000	3358.35	5536.00
	13	813.26	555.03	0.143	−275.57	1902.09
	14	−2121.87*	555.03	0.000	−3210.70	−1033.05
	15	3396.33*	555.03	0.000	2307.51	4485.16
	16	3011.47*	555.03	0.000	1922.64	4100.30
	17	195.17	555.03	0.725	−893.66	1284.00
	18	5491.89*	555.03	0.000	4403.06	6580.71
	19	1554.30*	555.03	0.005	465.47	2643.12
	20	−1288.71*	555.03	0.020	−2377.54	−199.88
	21	4066.31*	555.03	0.000	2977.48	5155.13
	22	466.26	555.03	0.401	−622.57	1555.09
	23	−2432.89*	555.03	0.000	−3521.72	−1344.06
	24	3027.00*	555.03	0.000	1938.17	4115.83
	25	2635.85*	555.03	0.000	1547.02	3724.68
	26	−151.93	555.03	0.784	−1240.76	936.90
	27	5099.88*	555.03	0.000	4011.05	6188.71
9	1	−3564.04*	555.03	0.000	−4652.87	−2475.22
	2	−6426.66*	555.03	0.000	−7515.49	−5337.83
	3	−1039.34	555.03	0.061	−2128.17	49.48
	4	−4659.60*	555.03	0.000	−5748.43	−3570.77
	5	−7580.28*	555.03	0.000	−8669.11	−6491.45
	6	−2084.63*	555.03	0.000	−3173.46	−995.80
	7	−2475.24*	555.03	0.000	−3564.06	−1386.41
	8	−5280.76*	555.03	0.000	−6369.59	−4191.93
	10	−3364.89*	555.03	0.000	−4453.72	−2276.06
	11	−6240.10*	555.03	0.000	−7328.93	−5151.27
	12	−833.58	555.03	0.133	−1922.41	255.25
	13	−4467.50*	555.03	0.000	−5556.33	−3378.67
	14	−7402.63*	555.03	0.000	−8491.46	−6313.80
	15	−1884.42*	555.03	0.001	−2973.25	−795.59
	16	−2269.29*	555.03	0.000	−3358.11	−1180.46
	17	−5085.59*	555.03	0.000	−6174.42	−3996.76
	18	211.13	555.03	0.704	−877.70	1299.96
	19	−3726.46*	555.03	0.000	−4815.29	−2637.63
	20	−6569.47*	555.03	0.000	−7658.30	−5480.64
	21	−1214.45*	555.03	0.029	−2303.28	−125.62
	22	−4814.50*	555.03	0.000	−5903.33	−3725.67
	23	−7713.65*	555.03	0.000	−8802.48	−6624.82
	24	−2253.76*	555.03	0.000	−3342.59	−1164.93
	25	−2644.91*	555.03	0.000	−3733.74	−1556.08
	26	−5432.69*	555.03	0.000	−6521.52	−4343.86
	27	−180.88	555.03	0.745	−1269.71	907.95

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
10	1	−199.15	555.03	0.720	−1287.98	889.68
	2	−3061.77*	555.03	0.000	−4150.60	−1972.94
	3	2325.55*	555.03	0.000	1236.72	3414.38
	4	−1294.71*	555.03	0.020	−2383.54	−205.88
	5	−4215.39*	555.03	0.000	−5304.22	−3126.56
	6	1280.27*	555.03	0.021	191.44	2369.09
	7	889.66	555.03	0.109	−199.17	1978.49
	8	−1915.87*	555.03	0.001	−3004.69	−827.04
	9	3364.89*	555.03	0.000	2276.06	4453.72
	11	−2875.20*	555.03	0.000	−3964.03	−1786.38
	12	2531.31*	555.03	0.000	1442.48	3620.14
	13	−1102.61*	555.03	0.047	−2191.43	−13.78
	14	−4037.74*	555.03	0.000	−5126.57	−2948.91
	15	1480.47*	555.03	0.008	391.64	2569.30
	16	1095.61*	555.03	0.049	6.78	2184.44
	17	−1720.70*	555.03	0.002	−2809.53	−631.87
	18	3576.02*	555.03	0.000	2487.19	4664.85
	19	−361.57	555.03	0.515	−1450.40	727.26
	20	−3204.57*	555.03	0.000	−4293.40	−2115.75
	21	2150.44*	555.03	0.000	1061.61	3239.27
	22	−1449.61*	555.03	0.009	−2538.43	−360.78
	23	−4348.75*	555.03	0.000	−5437.58	−3259.93
	24	1111.13*	555.03	0.045	22.31	2199.96
	25	719.98	555.03	0.195	−368.85	1808.81
	26	−2067.79*	555.03	0.000	−3156.62	−978.97
	27	3184.01*	555.03	0.000	2095.18	4272.84
11	1	2676.05*	555.03	0.000	1587.22	3764.88
	2	−186.56	555.03	0.737	−1275.39	902.26
	3	5200.75*	555.03	0.000	4111.92	6289.58
	4	1580.50*	555.03	0.004	491.67	2669.32
	5	−1340.19*	555.03	0.016	−2429.01	−251.36
	6	4155.47*	555.03	0.000	3066.64	5244.30
	7	3764.86*	555.03	0.000	2676.03	4853.69
	8	959.34	555.03	0.084	−129.49	2048.17
	9	6240.10*	555.03	0.000	5151.27	7328.93
	10	2875.20*	555.03	0.000	1786.38	3964.03
	12	5406.51*	555.03	0.000	4317.68	6495.34
	13	1772.60*	555.03	0.001	683.77	2861.43
	14	−1162.54*	555.03	0.036	−2251.36	−73.71
	15	4355.67*	555.03	0.000	3266.84	5444.50
	16	3970.81*	555.03	0.000	2881.98	5059.64
	17	1154.51*	555.03	0.038	65.68	2243.34
	18	6451.23*	555.03	0.000	5362.40	7540.05
	19	2513.63*	555.03	0.000	1424.81	3602.46
	20	−329.37	555.03	0.553	−1418.20	759.46
	21	5025.65*	555.03	0.000	3936.82	6114.47
	22	1425.60*	555.03	0.010	336.77	2514.43
	23	−1473.55*	555.03	0.008	−2562.38	−384.72
	24	3986.34*	555.03	0.000	2897.51	5075.17
	25	3595.19*	555.03	0.000	2506.36	4684.02
	26	807.41	555.03	0.146	−281.42	1896.24
	27	6059.22*	555.03	0.000	4970.39	7148.04

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
12	1	−2730.46*	555.03	0.000	−3819.29	−1641.63
	2	−5593.08*	555.03	0.000	−6681.91	−4504.25
	3	−205.76	555.03	0.711	−1294.59	883.07
	4	−3826.02*	555.03	0.000	−4914.85	−2737.19
	5	−6746.70*	555.03	0.000	−7835.53	−5657.87
	6	−1251.04*	555.03	0.024	−2339.87	−162.21
	7	−1641.65*	555.03	0.003	−2730.48	−552.82
	8	−4447.17*	555.03	0.000	−5536.00	−3358.35
	9	833.58	555.03	0.133	−255.25	1922.41
	10	−2531.31*	555.03	0.000	−3620.14	−1442.48
	11	−5406.51*	555.03	0.000	−6495.34	−4317.68
	13	−3633.91*	555.03	0.000	−4722.74	−2545.09
	14	−6569.05*	555.03	0.000	−7657.88	−5480.22
	15	−1050.84	555.03	0.059	−2139.67	37.99
	16	−1435.70*	555.03	0.010	−2524.53	−346.87
	17	−4252.01*	555.03	0.000	−5340.83	−3163.18
	18	1044.71	555.03	0.060	−44.12	2133.54
	19	−2892.88*	555.03	0.000	−3981.71	−1804.05
	20	−5735.88*	555.03	0.000	−6824.71	−4647.06
	21	−380.87	555.03	0.493	−1469.70	707.96
	22	−3980.91*	555.03	0.000	−5069.74	−2892.09
	23	−6880.06*	555.03	0.000	−7968.89	−5791.24
	24	−1420.18*	555.03	0.011	−2509.00	−331.35
	25	−1811.33*	555.03	0.001	−2900.15	−722.50
	26	−4599.10*	555.03	0.000	−5687.93	−3510.28
	27	652.70	555.03	0.240	−436.13	1741.53
13	1	903.45	555.03	0.104	−185.37	1992.28
	2	−1959.16*	555.03	0.000	−3047.99	−870.33
	3	3428.15*	555.03	0.000	2339.33	4516.98
	4	−192.10	555.03	0.729	−1280.93	896.72
	5	−3112.78*	555.03	0.000	−4201.61	−2023.96
	6	2382.87*	555.03	0.000	1294.04	3471.70
	7	1992.26*	555.03	0.000	903.43	3081.09
	8	−813.26	555.03	0.143	−1902.09	275.57
	9	4467.50*	555.03	0.000	3378.67	5556.33
	10	1102.61*	555.03	0.047	13.78	2191.43
	11	−1772.60*	555.03	0.001	−2861.43	−683.77
	12	3633.91*	555.03	0.000	2545.09	4722.74
	14	−2935.13*	555.03	0.000	−4023.96	−1846.31
	15	2583.07*	555.03	0.000	1494.25	3671.90
	16	2198.21*	555.03	0.000	1109.38	3287.04
	17	−618.09	555.03	0.266	−1706.92	470.74
	18	4678.63*	555.03	0.000	3589.80	5767.45
	19	741.04	555.03	0.182	−347.79	1829.86
	20	−2101.97*	555.03	0.000	−3190.80	−1013.14
	21	3253.05*	555.03	0.000	2164.22	4341.87
	22	−347.00	555.03	0.532	−1435.83	741.83
	23	−3246.15*	555.03	0.000	−4334.98	−2157.32
	24	2213.74*	555.03	0.000	1124.91	3302.57
	25	1822.59*	555.03	0.001	733.76	2911.42
	26	−965.19	555.03	0.082	−2054.02	123.64
	27	4286.62*	555.03	0.000	3197.79	5375.45

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
	1	3838.59*	555.03	0.000	2749.76	4927.42
	2	975.97	555.03	0.079	−112.86	2064.80
	3	6363.29*	555.03	0.000	5274.46	7452.12
	4	2743.03*	555.03	0.000	1654.20	3831.86
	5	−177.65	555.03	0.749	−1266.48	911.18
	6	5318.01*	555.03	0.000	4229.18	6406.83
	7	4927.40*	555.03	0.000	3838.57	6016.23
	8	2121.87*	555.03	0.000	1033.05	3210.70
	9	7402.63*	555.03	0.000	6313.80	8491.46
	10	4037.74*	555.03	0.000	2948.91	5126.57
	11	1162.54*	555.03	0.036	73.71	2251.36
	12	6569.05*	555.03	0.000	5480.22	7657.88
14	13	2935.13*	555.03	0.000	1846.31	4023.96
	15	5518.21*	555.03	0.000	4429.38	6607.04
	16	5133.35*	555.03	0.000	4044.52	6222.18
	17	2317.04*	555.03	0.000	1228.21	3405.87
	18	7613.76*	555.03	0.000	6524.93	8702.59
	19	3676.17*	555.03	0.000	2587.34	4765.00
	20	833.17	555.03	0.134	−255.66	1921.99
	21	6188.18*	555.03	0.000	5099.35	7277.01
	22	2588.13*	555.03	0.000	1499.31	3676.96
	23	−311.01	555.03	0.575	−1399.84	777.81
	24	5148.87*	555.03	0.000	4060.05	6237.70
	25	4757.72*	555.03	0.000	3668.89	5846.55
	26	1969.95*	555.03	0.000	881.12	3058.77
	27	7221.75*	555.03	0.000	6132.92	8310.58
	1	−1679.62*	555.03	0.003	−2768.45	−590.79
	2	−4542.24*	555.03	0.000	−5631.07	−3453.41
	3	845.08	555.03	0.128	−243.75	1933.91
	4	−2775.18*	555.03	0.000	−3864.01	−1686.35
	5	−5695.86*	555.03	0.000	−6784.69	−4607.03
	6	−200.20	555.03	0.718	−1289.03	888.63
	7	−590.81	555.03	0.287	−1679.64	498.02
	8	−3396.33*	555.03	0.000	−4485.16	−2307.51
	9	1884.42*	555.03	0.001	795.59	2973.25
	10	−1480.47*	555.03	0.008	−2569.30	−391.64
	11	−4355.67*	555.03	0.000	−5444.50	−3266.84
	12	1050.84	555.03	0.059	−37.99	2139.67
15	13	−2583.07*	555.03	0.000	−3671.90	−1494.25
	14	−5518.21*	555.03	0.000	−6607.04	−4429.38
	16	−384.86	555.03	0.488	−1473.69	703.97
	17	−3201.17*	555.03	0.000	−4289.99	−2112.34
	18	2095.55*	555.03	0.000	1006.72	3184.38
	19	−1842.04*	555.03	0.001	−2930.87	−753.21
	20	−4685.04*	555.03	0.000	−5773.87	−3596.22
	21	669.97	555.03	0.228	−418.86	1758.80
	22	−2930.07*	555.03	0.000	−4018.90	−1841.25
	23	−5829.22*	555.03	0.000	−6918.05	−4740.40
	24	−369.34	555.03	0.506	−1458.16	719.49
	25	−760.49	555.03	0.171	−1849.31	328.34
	26	−3548.26*	555.03	0.000	−4637.09	−2459.44
	27	1703.54*	555.03	0.002	614.71	2792.37

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
16	1	−1294.76*	555.03	0.020	−2383.59	−205.93
	2	−4157.38*	555.03	0.000	−5246.20	−3068.55
	3	1229.94*	555.03	0.027	141.11	2318.77
	4	−2390.32*	555.03	0.000	−3479.15	−1301.49
	5	−5311.00*	555.03	0.000	−6399.83	−4222.17
	6	184.66	555.03	0.739	−904.17	1273.49
	7	−205.95	555.03	0.711	−1294.78	882.88
	8	−3011.47*	555.03	0.000	−4100.30	−1922.64
	9	2269.29*	555.03	0.000	1180.46	3358.11
	10	−1095.61*	555.03	0.049	−2184.44	−6.78
	11	−3970.81*	555.03	0.000	−5059.64	−2881.98
	12	1435.70*	555.03	0.010	346.87	2524.53
	13	−2198.21*	555.03	0.000	−3287.04	−1109.38
	14	−5133.35*	555.03	0.000	−6222.18	−4044.52
	15	384.86	555.03	0.488	−703.97	1473.69
	17	−2816.30*	555.03	0.000	−3905.13	−1727.48
	18	2480.41*	555.03	0.000	1391.58	3569.24
	19	−1457.18*	555.03	0.009	−2546.01	−368.35
	20	−4300.18*	555.03	0.000	−5389.01	−3211.35
	21	1054.83	555.03	0.058	−34.00	2143.66
	22	−2545.21*	555.03	0.000	−3634.04	−1456.38
	23	−5444.36*	555.03	0.000	−6533.19	−4355.53
	24	15.53	555.03	0.978	−1073.30	1104.36
	25	−375.62	555.03	0.499	−1464.45	713.20
	26	−3163.40*	555.03	0.000	−4252.23	−2074.57
	27	2088.40*	555.03	0.000	999.58	3177.23
17	1	1521.55*	555.03	0.006	432.72	2610.37
	2	−1341.07*	555.03	0.016	−2429.90	−252.24
	3	4046.25*	555.03	0.000	2957.42	5135.07
	4	425.99	555.03	0.443	−662.84	1514.82
	5	−2494.69*	555.03	0.000	−3583.52	−1405.86
	6	3000.96*	555.03	0.000	1912.13	4089.79
	7	2610.35*	555.03	0.000	1521.52	3699.18
	8	−195.17	555.03	0.725	−1284.00	893.66
	9	5085.59*	555.03	0.000	3996.76	6174.42
	10	1720.70*	555.03	0.002	631.87	2809.53
	11	−1154.51*	555.03	0.038	−2243.34	−65.68
	12	4252.01*	555.03	0.000	3163.18	5340.83
	13	618.09	555.03	0.266	−470.74	1706.92
	14	−2317.04*	555.03	0.000	−3405.87	−1228.21
	15	3201.17*	555.03	0.000	2112.34	4289.99
	16	2816.30*	555.03	0.000	1727.48	3905.13
	18	5296.72*	555.03	0.000	4207.89	6385.55
	19	1359.13*	555.03	0.014	270.30	2447.96
	20	−1483.88*	555.03	0.008	−2572.71	−395.05
	21	3871.14*	555.03	0.000	2782.31	4959.97
	22	271.09	555.03	0.625	−817.74	1359.92
	23	−2628.06*	555.03	0.000	−3716.89	−1539.23
	24	2831.83*	555.03	0.000	1743.00	3920.66
	25	2440.68*	555.03	0.000	1351.85	3529.51
	26	−347.10	555.03	0.532	−1435.93	741.73
	27	4904.71*	555.03	0.000	3815.88	5993.54

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
	1	−3775.17*	555.03	0.000	−4864.00	−2686.34
	2	−6637.79*	555.03	0.000	−7726.62	−5548.96
	3	−1250.47*	555.03	0.024	−2339.30	−161.64
	4	−4870.73*	555.03	0.000	−5959.56	−3781.90
	5	−7791.41*	555.03	0.000	−8880.24	−6702.58
	6	−2295.76*	555.03	0.000	−3384.58	−1206.93
	7	−2686.36*	555.03	0.000	−3775.19	−1597.54
	8	−5491.89*	555.03	0.000	−6580.71	−4403.06
	9	−211.13	555.03	0.704	−1299.96	877.70
	10	−3576.02*	555.03	0.000	−4664.85	−2487.19
	11	−6451.23*	555.03	0.000	−7540.05	−5362.40
	12	−1044.71	555.03	0.060	−2133.54	44.12
18	13	−4678.63*	555.03	0.000	−5767.45	−3589.80
	14	−7613.76*	555.03	0.000	−8702.59	−6524.93
	15	−2095.55*	555.03	0.000	−3184.38	−1006.72
	16	−2480.41*	555.03	0.000	−3569.24	−1391.58
	17	−5296.72*	555.03	0.000	−6385.55	−4207.89
	19	−3937.59*	555.03	0.000	−5026.42	−2848.76
	20	−6780.60*	555.03	0.000	−7869.42	−5691.77
	21	−1425.58*	555.03	0.010	−2514.41	−336.75
	22	−5025.63*	555.03	0.000	−6114.45	−3936.80
	23	−7924.78*	555.03	0.000	−9013.60	−6835.95
	24	−2464.89*	555.03	0.000	−3553.72	−1376.06
	25	−2856.04*	555.03	0.000	−3944.87	−1767.21
	26	−5643.82*	555.03	0.000	−6732.64	−4554.99
	27	−392.01	555.03	0.480	−1480.84	696.82
	1	162.42	555.03	0.770	−926.41	1251.25
	2	−2700.20*	555.03	0.000	−3789.03	−1611.37
	3	2687.12*	555.03	0.000	1598.29	3775.95
	4	−933.14	555.03	0.093	−2021.97	155.69
	5	−3853.82*	555.03	0.000	−4942.65	−2764.99
	6	1641.84*	555.03	0.003	553.01	2730.66
	7	1251.23*	555.03	0.024	162.40	2340.06
	8	−1554.30*	555.03	0.005	−2643.12	−465.47
	9	3726.46*	555.03	0.000	2637.63	4815.29
	10	361.57	555.03	0.515	−727.26	1450.40
	11	−2513.63*	555.03	0.000	−3602.46	−1424.81
	12	2892.88*	555.03	0.000	1804.05	3981.71
19	13	−741.04	555.03	0.182	−1829.86	347.79
	14	−3676.17*	555.03	0.000	−4765.00	−2587.34
	15	1842.04*	555.03	0.001	753.21	2930.87
	16	1457.18*	555.03	0.009	368.35	2546.01
	17	−1359.13*	555.03	0.014	−2447.96	−270.30
	18	3937.59*	555.03	0.000	2848.76	5026.42
	20	−2843.00*	555.03	0.000	−3931.83	−1754.18
	21	2512.01*	555.03	0.000	1423.18	3600.84
	22	−1088.04	555.03	0.050	−2176.86	.79
	23	−3987.18*	555.03	0.000	−5076.01	−2898.36
	24	1472.70*	555.03	0.008	383.88	2561.53
	25	1081.55	555.03	0.052	−7.28	2170.38
	26	−1706.22*	555.03	0.002	−2795.05	−617.40
	27	3545.58*	555.03	0.000	2456.75	4634.41

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
20	1	3005.42*	555.03	0.000	1916.59	4094.25
	2	142.81	555.03	0.797	−946.02	1231.63
	3	5530.12*	555.03	0.000	4441.29	6618.95
	4	1909.87*	555.03	0.001	821.04	2998.69
	5	−1010.82	555.03	0.069	−2099.64	78.01
	6	4484.84*	555.03	0.000	3396.01	5573.67
	7	4094.23*	555.03	0.000	3005.40	5183.06
	8	1288.71*	555.03	0.020	199.88	2377.54
	9	6569.47*	555.03	0.000	5480.64	7658.30
	10	3204.57*	555.03	0.000	2115.75	4293.40
	11	329.37	555.03	0.553	−759.46	1418.20
	12	5735.88*	555.03	0.000	4647.06	6824.71
	13	2101.97*	555.03	0.000	1013.14	3190.80
	14	−833.17	555.03	0.134	−1921.99	255.66
	15	4685.04*	555.03	0.000	3596.22	5773.87
	16	4300.18*	555.03	0.000	3211.35	5389.01
	17	1483.88*	555.03	0.008	395.05	2572.71
	18	6780.60*	555.03	0.000	5691.77	7869.42
	19	2843.00*	555.03	0.000	1754.18	3931.83
	21	5355.02*	555.03	0.000	4266.19	6443.84
	22	1754.97*	555.03	0.002	666.14	2843.80
	23	−1144.18*	555.03	0.039	−2233.01	−55.35
	24	4315.71*	555.03	0.000	3226.88	5404.54
	25	3924.56*	555.03	0.000	2835.73	5013.39
	26	1136.78*	555.03	0.041	47.95	2225.61
	27	6388.59*	555.03	0.000	5299.76	7477.41
21	1	−2349.59*	555.03	0.000	−3438.42	−1260.76
	2	−5212.21*	555.03	0.000	−6301.04	−4123.38
	3	175.11	555.03	0.752	−913.72	1263.94
	4	−3445.15*	555.03	0.000	−4533.98	−2356.32
	5	−6365.83*	555.03	0.000	−7454.66	−5277.00
	6	−870.18	555.03	0.117	−1959.00	218.65
	7	−1260.78*	555.03	0.023	−2349.61	−171.96
	8	−4066.31*	555.03	0.000	−5155.13	−2977.48
	9	1214.45*	555.03	0.029	125.62	2303.28
	10	−2150.44*	555.03	0.000	−3239.27	−1061.61
	11	−5025.65*	555.03	0.000	−6114.47	−3936.82
	12	380.87	555.03	0.493	−707.96	1469.70
	13	−3253.05*	555.03	0.000	−4341.87	−2164.22
	14	−6188.18*	555.03	0.000	−7277.01	−5099.35
	15	−669.97	555.03	0.228	−1758.80	418.86
	16	−1054.83	555.03	0.058	−2143.66	34.00
	17	−3871.14*	555.03	0.000	−4959.97	−2782.31
	18	1425.58*	555.03	0.010	336.75	2514.41
	19	−2512.01*	555.03	0.000	−3600.84	−1423.18
	20	−5355.02*	555.03	0.000	−6443.84	−4266.19
	22	−3600.05*	555.03	0.000	−4688.87	−2511.22
	23	−6499.20*	555.03	0.000	−7588.02	−5410.37
	24	−1039.31	555.03	0.061	−2128.14	49.52
	25	−1430.46*	555.03	0.010	−2519.29	−341.63
	26	−4218.24*	555.03	0.000	−5307.06	−3129.41
	27	1033.57	555.03	0.063	−55.26	2122.40

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
	1	1250.45*	555.03	0.024	161.63	2339.28
	2	−1612.16*	555.03	0.004	−2700.99	−523.33
	3	3775.15*	555.03	0.000	2686.33	4863.98
	4	154.90	555.03	0.780	−933.93	1243.72
	5	−2765.78*	555.03	0.000	−3854.61	−1676.96
	6	2729.87*	555.03	0.000	1641.04	3818.70
	7	2339.26*	555.03	0.000	1250.43	3428.09
	8	−466.26	555.03	0.401	−1555.09	622.57
	9	4814.50*	555.03	0.000	3725.67	5903.33
	10	1449.61*	555.03	0.009	360.78	2538.43
	11	−1425.60*	555.03	0.010	−2514.43	−336.77
	12	3980.91*	555.03	0.000	2892.09	5069.74
22	13	347.00	555.03	0.532	−741.83	1435.83
	14	−2588.13*	555.03	0.000	−3676.96	−1499.31
	15	2930.07*	555.03	0.000	1841.25	4018.90
	16	2545.21*	555.03	0.000	1456.38	3634.04
	17	−271.09	555.03	0.625	−1359.92	817.74
	18	5025.63*	555.03	0.000	3936.80	6114.45
	19	1088.04	555.03	0.050	−0.79	2176.86
	20	−1754.97*	555.03	0.002	−2843.80	−666.14
	21	3600.05*	555.03	0.000	2511.22	4688.87
	23	−2899.15*	555.03	0.000	−3987.98	−1810.32
	24	2560.74*	555.03	0.000	1471.91	3649.57
	25	2169.59*	555.03	0.000	1080.76	3258.42
	26	−618.19	555.03	0.266	−1707.02	470.64
	27	4633.62*	555.03	0.000	3544.79	5722.45
	1	4149.60*	555.03	0.000	3060.77	5238.43
	2	1286.99*	555.03	0.021	198.16	2375.81
	3	6674.30*	555.03	0.000	5585.47	7763.13
	4	3054.05*	555.03	0.000	1965.22	4142.87
	5	133.36	555.03	0.810	−955.46	1222.19
	6	5629.02*	555.03	0.000	4540.19	6717.85
	7	5238.41*	555.03	0.000	4149.58	6327.24
	8	2432.89*	555.03	0.000	1344.06	3521.72
	9	7713.65*	555.03	0.000	6624.82	8802.48
	10	4348.75*	555.03	0.000	3259.93	5437.58
	11	1473.55*	555.03	0.008	384.72	2562.38
	12	6880.06*	555.03	0.000	5791.24	7968.89
23	13	3246.15*	555.03	0.000	2157.32	4334.98
	14	311.01	555.03	0.575	−777.81	1399.84
	15	5829.22*	555.03	0.000	4740.40	6918.05
	16	5444.36*	555.03	0.000	4355.53	6533.19
	17	2628.06*	555.03	0.000	1539.23	3716.89
	18	7924.78*	555.03	0.000	6835.95	9013.60
	19	3987.18*	555.03	0.000	2898.36	5076.01
	20	1144.18*	555.03	0.039	55.35	2233.01
	21	6499.20*	555.03	0.000	5410.37	7588.02
	22	2899.15*	555.03	0.000	1810.32	3987.98
	24	5459.89*	555.03	0.000	4371.06	6548.72
	25	5068.74*	555.03	0.000	3979.91	6157.57
	26	2280.96*	555.03	0.000	1192.13	3369.79
	27	7532.77*	555.03	0.000	6443.94	8621.59

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
24	1	−1310.29*	555.03	0.018	−2399.11	−221.46
	2	−4172.90*	555.03	0.000	−5261.73	−3084.07
	3	1214.41*	555.03	0.029	125.59	2303.24
	4	−2405.84*	555.03	0.000	−3494.67	−1317.01
	5	−5326.52*	555.03	0.000	−6415.35	−4237.70
	6	169.13	555.03	0.761	−919.70	1257.96
	7	−221.48	555.03	0.690	−1310.31	867.35
	8	−3027.00*	555.03	0.000	−4115.83	−1938.17
	9	2253.76*	555.03	0.000	1164.93	3342.59
	10	−1111.13*	555.03	0.045	−2199.96	−22.31
	11	−3986.34*	555.03	0.000	−5075.17	−2897.51
	12	1420.18*	555.03	0.011	331.35	2509.00
	13	−2213.74*	555.03	0.000	−3302.57	−1124.91
	14	−5148.87*	555.03	0.000	−6237.70	−4060.05
	15	369.34	555.03	0.506	−719.49	1458.16
	16	−15.53	555.03	0.978	−1104.36	1073.30
	17	−2831.83*	555.03	0.000	−3920.66	−1743.00
	18	2464.89*	555.03	0.000	1376.06	3553.72
	19	−1472.70*	555.03	0.008	−2561.53	−383.88
	20	−4315.71*	555.03	0.000	−5404.54	−3226.88
	21	1039.31	555.03	0.061	−49.52	2128.14
	22	−2560.74*	555.03	0.000	−3649.57	−1471.91
	23	−5459.89*	555.03	0.000	−6548.72	−4371.06
	25	−391.15	555.03	0.481	−1479.98	697.68
	26	−3178.93*	555.03	0.000	−4267.76	−2090.10
	27	2072.88*	555.03	0.000	984.05	3161.71
25	1	−919.13	555.03	0.098	−2007.96	169.69
	2	−3781.75*	555.03	0.000	−4870.58	−2692.92
	3	1605.57*	555.03	0.004	516.74	2694.39
	4	−2014.69*	555.03	0.000	−3103.52	−925.86
	5	−4935.37*	555.03	0.000	−6024.20	−3846.54
	6	560.28	555.03	0.313	−528.55	1649.11
	7	169.67	555.03	0.760	−919.16	1258.50
	8	−2635.85*	555.03	0.000	−3724.68	−1547.02
	9	2644.91*	555.03	0.000	1556.08	3733.74
	10	−719.98	555.03	0.195	−1808.81	368.85
	11	−3595.19*	555.03	0.000	−4684.02	−2506.36
	12	1811.33*	555.03	0.001	722.50	2900.15
	13	−1822.59*	555.03	0.001	−2911.42	−733.76
	14	−4757.72*	555.03	0.000	−5846.55	−3668.89
	15	760.49	555.03	0.171	−328.34	1849.31
	16	375.62	555.03	0.499	−713.20	1464.45
	17	−2440.68*	555.03	0.000	−3529.51	−1351.85
	18	2856.04*	555.03	0.000	1767.21	3944.87
	19	−1081.55	555.03	0.052	−2170.38	7.28
	20	−3924.56*	555.03	0.000	−5013.39	−2835.73
	21	1430.46*	555.03	0.010	341.63	2519.29
	22	−2169.59*	555.03	0.000	−3258.42	−1080.76
	23	−5068.74*	555.03	0.000	−6157.57	−3979.91
	24	391.15	555.03	0.481	−697.68	1479.98
	26	−2787.78*	555.03	0.000	−3876.61	−1698.95
	27	2464.03*	555.03	0.000	1375.20	3552.86

TABLE 6: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. Error	Sig.	95% confidence interval	
					Lower bound	Upper bound
26	1	1868.64*	555.03	0.001	779.81	2957.47
	2	−993.97	555.03	0.074	−2082.80	94.86
	3	4393.34*	555.03	0.000	3304.51	5482.17
	4	773.09	555.03	0.164	−315.74	1861.91
	5	−2147.60*	555.03	0.000	−3236.42	−1058.77
	6	3348.06*	555.03	0.000	2259.23	4436.89
	7	2957.45*	555.03	0.000	1868.62	4046.28
	8	151.93	555.03	0.784	−936.90	1240.76
	9	5432.69*	555.03	0.000	4343.86	6521.52
	10	2067.79*	555.03	0.000	978.97	3156.62
	11	−807.41	555.03	0.146	−1896.24	281.42
	12	4599.10*	555.03	0.000	3510.28	5687.93
	13	965.19	555.03	0.082	−123.64	2054.02
	14	−1969.95*	555.03	0.000	−3058.77	−881.12
	15	3548.26*	555.03	0.000	2459.44	4637.09
	16	3163.40*	555.03	0.000	2074.57	4252.23
	17	347.10	555.03	0.532	−741.73	1435.93
	18	5643.82*	555.03	0.000	4554.99	6732.64
	19	1706.22*	555.03	0.002	617.40	2795.05
	20	−1136.78*	555.03	0.041	−2225.61	−47.95
	21	4218.24*	555.03	0.000	3129.41	5307.06
	22	618.19	555.03	0.266	−470.64	1707.02
	23	−2280.96*	555.03	0.000	−3369.79	−1192.13
	24	3178.93*	555.03	0.000	2090.10	4267.76
	25	2787.78*	555.03	0.000	1698.95	3876.61
	27	5251.81*	555.03	0.000	4162.98	6340.63
27	1	−3383.16*	555.03	0.000	−4471.99	−2294.33
	2	−6245.78*	555.03	0.000	−7334.61	−5156.95
	3	−858.46	555.03	0.122	−1947.29	230.37
	4	−4478.72*	555.03	0.000	−5567.55	−3389.89
	5	−7399.40*	555.03	0.000	−8488.23	−6310.57
	6	−1903.75*	555.03	0.001	−2992.57	−814.92
	7	−2294.35*	555.03	0.000	−3383.18	−1205.53
	8	−5099.88*	555.03	0.000	−6188.71	−4011.05
	9	180.88	555.03	0.745	−907.95	1269.71
	10	−3184.01*	555.03	0.000	−4272.84	−2095.18
	11	−6059.22*	555.03	0.000	−7148.04	−4970.39
	12	−652.70	555.03	0.240	−1741.53	436.13
	13	−4286.62*	555.03	0.000	−5375.45	−3197.79
	14	−7221.75*	555.03	0.000	−8310.58	−6132.92
	15	−1703.54*	555.03	0.002	−2792.37	−614.71
	16	−2088.40*	555.03	0.000	−3177.23	−999.58
	17	−4904.71*	555.03	0.000	−5993.54	−3815.88
	18	392.01	555.03	0.480	−696.82	1480.84
	19	−3545.58*	555.03	0.000	−4634.41	−2456.75
	20	−6388.59*	555.03	0.000	−7477.41	−5299.76
	21	−1033.57	555.03	0.063	−2122.40	55.26
	22	−4633.62*	555.03	0.000	−5722.45	−3544.79
	23	−7532.77*	555.03	0.000	−8621.59	−6443.94
	24	−2072.88*	555.03	0.000	−3161.71	−984.05
	25	−2464.03*	555.03	0.000	−3552.86	−1375.20
	26	−5251.81*	555.03	0.000	−6340.63	−4162.98

*The mean difference is significant at the 0.05 level.

TABLE 7: Welch's ANOVA.

Annual_Patent_Application_Acceptance			
Statistic[a]	df1	df2	Sig.
Welch	141.386	26	474.820	0.000

[a]Asymptotically F distributed.

TABLE 8: Multiple comparisons on dependent variable Annual_Patent_Application_Acceptance.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
1	2	−1334561.15*	130309.35	0.000	−1831489.21	−837633.09
	3	1035494.80*	102289.44	0.000	645497.88	1425491.72
	4	−279970.62	115257.21	0.773	−718756.81	158815.57
	5	−1677613.28*	133154.93	0.000	−2185637.58	−1169588.98
	6	803578.50*	103540.25	0.000	408957.30	1198199.71
	7	272330.42	111560.80	0.766	−152383.64	697044.47
	8	−1001451.13*	127580.66	0.000	−1487763.65	−515138.61
	9	1261452.03*	101091.25	0.000	875863.92	1647040.14
	10	−5183.39	114421.05	1.000	−440777.19	430410.41
	11	−1367536.94*	132259.72	0.000	−1872067.76	−863006.12
	12	1046576.35*	102840.85	0.000	654543.52	1438609.19
	13	−292030.25	116426.81	0.720	−735290.17	151229.67
	14	−1720718.86*	135305.26	0.000	−2237143.77	−1204293.95
	15	809955.73*	104160.56	0.000	413033.36	1206878.10
	16	273579.28	112501.66	0.772	−154706.32	701864.87
	17	−1024941.70*	129343.67	0.000	−1518109.97	−531773.43
	18	1276929.24*	101578.27	0.000	889551.66	1664306.82
	19	19204.88	112248.20	1.000	−408117.84	446527.60
	20	−1283466.38*	128269.37	0.000	−1772455.73	−794477.03
	21	1035215.93*	101686.92	0.000	647438.68	1422993.18
	22	−253584.42	114019.61	0.885	−687647.34	180478.50
	23	−1616054.61*	130920.72	0.000	−2115364.60	−1116744.62
	24	808265.88*	102869.27	0.000	416127.98	1200403.79
	25	284792.10	110548.05	0.672	−136086.08	705670.28
	26	−960192.23*	125723.50	0.000	−1439295.57	−481088.89
	27	1256511.66*	100552.72	0.000	872898.01	1640125.31

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
2	1	1334561.15*	130309.35	0.000	837633.09	1831489.21
	3	2370055.95*	120793.10	0.000	1907536.83	2832575.07
	4	1054590.53*	131954.79	0.000	551592.83	1557588.23
	5	−343052.13	147844.77	0.838	−905901.73	219797.47
	6	2138139.66*	121854.12	0.000	1671852.25	2604427.06
	7	1606891.57*	128738.71	0.000	1115728.06	2098055.07
	8	333110.02	142844.91	0.832	−210705.14	876925.18
	9	2596013.18*	119780.14	0.000	2137072.98	3054953.38
	10	1329377.76*	131225.07	0.000	829075.53	1829679.99
	11	−32975.79	147039.03	1.000	−592747.86	526796.28
	12	2381137.51*	121260.39	0.000	1916961.17	2845313.84
	13	1042530.90*	132977.61	0.000	535745.48	1549316.32
	14	−386157.71	149784.37	0.670	−956429.94	184114.52
	15	2144516.88*	122381.64	0.000	1676348.80	2612684.96
	16	1608140.43*	129554.87	0.000	1113985.11	2102295.74
	17	309619.45	144421.70	0.918	−240181.90	859420.80
	18	2611490.39*	120191.47	0.000	2151099.19	3071881.59
	19	1353766.03*	129334.84	0.000	860418.09	1847113.97
	20	51094.77	143460.36	1.000	−495054.97	597244.51
	21	2369777.08*	120283.30	0.000	1909061.50	2830492.66
	22	1080976.73*	130875.18	0.000	581964.86	1579988.60
	23	−281493.46	145835.79	0.972	−836676.72	273689.80
	24	2142827.04*	121284.50	0.000	1678565.09	2607088.98
	25	1619353.25*	127862.10	0.000	1131393.93	2107312.57
	26	374368.92	141188.68	0.616	−163176.58	911914.42
	27	2591072.81*	119325.99	0.000	2133731.03	3048414.59
3	1	−1035494.80*	102289.44	0.000	−1425491.72	−645497.88
	2	−2370055.95*	120793.10	0.000	−2832575.07	−1907536.83
	4	−1315465.42*	104377.54	0.000	−1713595.91	−917334.93
	5	−2713108.08*	123857.50	0.000	−3187702.83	−2238513.33
	6	−231916.29	91274.47	0.697	−579398.77	115566.18
	7	−763164.38*	100280.89	0.000	−1145355.72	−380973.05
	8	−2036945.93*	117844.27	0.000	−2487856.77	−1586035.09
	9	225957.23	88486.65	0.688	−110912.12	562826.58
	10	−1040678.19*	103453.49	0.000	−1435207.09	−646149.29
	11	−2403031.74*	122894.59	0.000	−2873830.79	−1932232.69
	12	11081.55	90480.31	1.000	−333369.89	355533.00
	13	−1327525.05*	105667.63	0.000	−1730689.26	−924360.84
	14	−2756213.66*	126166.40	0.000	−3239913.82	−2272513.50
	15	−225539.07	91977.54	0.758	−575709.66	124631.52
	16	−761915.52*	101326.55	0.000	−1148168.10	−375662.95
	17	−2060436.50*	119750.71	0.000	−2518850.71	−1602022.29
	18	241434.44	89042.64	0.570	−97545.33	580414.21
	19	−1016289.92*	101045.06	0.000	−1401448.69	−631131.15
	20	−2318961.18*	118589.54	0.000	−2772804.59	−1865117.77
	21	−278.87	89166.56	1.000	−339729.45	339171.71
	22	−1289079.22*	103009.31	0.000	−1681878.12	−896280.32
	23	−2651549.41*	121452.39	0.000	−3116665.54	−2186433.28
	24	−227228.92	90512.62	0.719	−571803.55	117345.72
	25	−750702.70*	99153.01	0.000	−1128519.54	−372885.86
	26	−1995687.03*	115831.11	0.000	−2438680.97	−1552693.09
	27	221016.86	87870.91	0.716	−113519.20	555552.92

Table 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
4	1	279970.62	115257.21	0.773	−158815.57	718756.81
	2	−1054590.53*	131954.79	0.000	−1557588.23	−551592.83
	3	1315465.42*	104377.54	0.000	917334.93	1713595.91
	5	−1397642.66*	134765.63	0.000	−1911576.98	−883708.34
	6	1083549.13*	105603.62	0.000	680908.27	1486189.98
	7	552301.04*	113478.42	0.001	120246.68	984355.39
	8	−721480.51*	129260.84	0.000	−1214023.15	−228937.87
	9	1541422.65*	103203.59	0.000	1147589.85	1935255.45
	10	274787.23	116291.52	0.814	−167925.62	717500.08
	11	−1087566.32*	133881.19	0.000	−1598056.55	−577076.09
	12	1326546.97*	104917.97	0.000	926431.06	1726662.89
	13	−12059.63	118265.56	1.000	−462289.77	438170.51
	14	−1440748.24*	136890.67	0.000	−1962967.51	−918528.97
	15	1089926.35*	106211.88	0.000	685040.03	1494812.67
	16	553549.90*	114403.51	0.001	117997.97	989101.82
	17	−744971.08*	131001.24	0.000	−1244264.94	−245677.22
	18	1556899.86*	103680.70	0.000	1161322.99	1952476.73
	19	299175.50	114154.28	0.639	−135433.36	733784.36
	20	−1003495.76*	129940.65	0.000	−1498674.02	−508317.50
	21	1315186.55*	103787.14	0.000	919220.09	1711153.01
	22	26386.20	115896.56	1.000	−414826.18	467598.58
	23	−1336083.99*	132558.58	0.000	−1841428.67	−830739.31
	24	1088236.50*	104945.84	0.000	688018.11	1488454.90
	25	564762.72*	112482.95	0.001	136463.31	993062.13
	26	−680221.61*	127428.18	0.000	−1165668.65	−194774.57
	27	1536482.28*	102676.15	0.000	1144573.36	1928391.20
5	1	1677613.28*	133154.93	0.000	1169588.98	2185637.58
	2	343052.13	147844.77	0.838	−219797.47	905901.73
	3	2713108.08*	123857.50	0.000	2238513.33	3187702.83
	4	1397642.66*	134765.63	0.000	883708.34	1911576.98
	6	2481191.79*	124892.49	0.000	2002942.30	2959441.27
	7	1949943.70*	131618.24	0.000	1447528.78	2452358.61
	8	676162.15*	145445.44	0.003	122389.78	1229934.52
	9	2939065.31*	122869.81	0.000	2467940.20	3410190.42
	10	1672429.89*	134051.22	0.000	1161120.62	2183739.16
	11	310076.34	149566.65	0.940	−259310.88	879463.56
	12	2724189.64*	124313.28	0.000	2247987.79	3200391.48
	13	1385583.03*	135767.28	0.000	867958.68	1903207.38
	14	−43105.58	152266.43	1.000	−622779.16	536568.00
	15	2487569.01*	125407.23	0.000	2007494.93	2967643.09
	16	1951192.56*	132416.67	0.000	1445866.80	2456518.31
	17	652671.58*	146994.34	0.006	93043.85	1212299.31
	18	2954542.52*	123270.83	0.000	2482010.88	3427074.16
	19	1696818.16*	132201.39	0.000	1192278.02	2201358.30
	20	394146.90	146049.93	0.580	−161908.66	950202.46
	21	2712829.21*	123360.37	0.000	2239983.10	3185675.32
	22	1424028.86*	133708.72	0.000	913975.98	1934081.74
	23	61558.67	148383.92	1.000	−503335.93	626453.27
	24	2485879.17*	124336.79	0.000	2009594.30	2962164.03
	25	1962405.38*	130760.94	0.000	1463106.86	2461703.90
	26	717421.05*	143819.16	0.001	169777.97	1265064.13
	27	2934124.94*	122427.13	0.000	2464549.00	3403700.88

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
6	1	−803578.50*	103540.25	0.000	−1198199.71	−408957.30
	2	−2138139.66*	121854.12	0.000	−2604427.06	−1671852.25
	3	231916.29	91274.47	0.697	−115566.18	579398.77
	4	−1083549.12*	105603.62	0.000	−1486189.98	−680908.27
	5	−2481191.78*	124892.49	0.000	−2959441.27	−2002942.30
	7	−531248.09*	101556.44	0.000	−918178.22	−144317.96
	8	−1805029.63*	118931.59	0.000	−2259824.58	−1350234.69
	9	457873.52*	89929.63	0.001	115482.10	800264.95
	10	−808761.89*	104690.38	0.000	−1207850.95	−409672.84
	11	−2171115.45*	123937.61	0.000	−2645604.20	−1696626.69
	12	242997.85	91892.00	0.621	−106828.81	592824.51
	13	−1095608.75*	106878.91	0.000	−1503215.45	−688002.06
	14	−2524297.37*	127182.60	0.000	−3011570.98	−2037023.75
	15	6377.22	93366.59	1.000	−349062.36	361816.81
	16	−529999.23*	102589.09	0.000	−920930.37	−139068.09
	17	−1828520.20*	120820.88	0.000	−2290742.88	−1366297.53
	18	473350.73*	90476.76	0.000	128890.47	817811.00
	19	−784373.62*	102311.08	0.000	−1174227.05	−394520.20
	20	−2087044.88*	119670.09	0.000	−2544742.49	−1629347.28
	21	231637.42	90598.72	0.686	−113284.44	576559.29
	22	−1057162.92*	104251.47	0.000	−1454546.26	−659779.59
	23	−2419633.12*	122507.69	0.000	−2888492.54	−1950773.69
	24	4687.38	91923.81	1.000	−345260.14	354634.90
	25	−518786.41*	100442.88	0.000	−901408.56	−136164.25
	26	−1763770.74*	116937.16	0.000	−2210731.81	−1316809.66
	27	452933.16*	89323.84	0.001	112828.38	793037.93
7	1	−272330.42	111560.80	0.766	−697044.47	152383.64
	2	−1606891.57*	128738.71	0.000	−2098055.07	−1115728.06
	3	763164.38*	100280.89	0.000	380973.05	1145355.72
	4	−552301.04*	113478.42	0.001	−984355.39	−120246.68
	5	−1949943.70*	131618.24	0.000	−2452358.61	−1447528.78
	6	531248.09*	101556.44	0.000	144317.96	918178.22
	8	−1273781.54*	125976.00	0.000	−1754173.15	−793389.94
	9	989121.61*	99058.41	0.000	611451.03	1366792.20
	10	−277513.80	112629.06	0.751	−706313.83	151286.22
	11	−1639867.36*	130712.51	0.000	−2138740.81	−1140993.90
	12	774245.94*	100843.28	0.000	389967.92	1158523.96
	13	−564360.66*	114666.17	0.001	−1000973.80	−127747.53
	14	−1993049.28*	133793.28	0.000	−2503977.02	−1482121.53
	15	537625.31*	102188.80	0.000	148338.12	926912.51
	16	1248.86	110678.59	1.000	−420097.15	422594.87
	17	−1297272.12*	127761.16	0.000	−1784621.36	−809922.87
	18	1004598.82*	99555.38	0.000	625093.01	1384104.64
	19	−253125.54	110420.95	0.853	−673489.03	167237.96
	20	−1555796.79*	126673.44	0.000	−2038905.42	−1072688.17
	21	762885.51*	99666.23	0.000	382969.87	1142801.16
	22	−525914.83*	112221.20	0.003	−953153.90	−98675.77
	23	−1888385.03*	129357.50	0.000	−2381964.45	−1394805.60
	24	535935.47*	100872.27	0.000	151549.78	920321.16
	25	12461.68	108692.23	1.000	−401323.17	426246.54
	26	−1232522.65*	124094.83	0.000	−1705594.20	−759451.09
	27	984181.25*	98508.77	0.000	608536.13	1359826.36

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
8	1	1001451.13*	127580.66	0.000	515138.61	1487763.65
	2	−333110.02	142844.91	0.832	−876925.18	210705.14
	3	2036945.93*	117844.27	0.000	1586035.09	2487856.77
	4	721480.51*	129260.84	0.000	228937.87	1214023.15
	5	−676162.15*	145445.44	0.003	−1229934.52	−122389.78
	6	1805029.63*	118931.59	0.000	1350234.69	2259824.58
	7	1273781.55*	125976.00	0.000	793389.94	1754173.15
	9	2262903.16*	116805.74	0.000	1815682.87	2710123.45
	10	996267.74*	128515.83	0.000	506491.31	1486044.17
	11	−366085.81	144626.32	0.704	−916717.97	184546.35
	12	2048027.48*	118323.21	0.000	1595408.26	2500646.71
	13	709420.88*	130304.81	0.000	212992.48	1205849.28
	14	−719267.73*	147416.60	0.001	−1280610.68	−157924.78
	15	1811406.86*	119472.02	0.000	1354674.09	2268139.63
	16	1275030.41*	126809.95	0.000	791565.35	1758495.46
	17	−23490.57	141964.52	1.000	−563943.32	516962.18
	18	2278380.37*	117227.50	0.000	1829663.62	2727097.12
	19	1020656.01*	126585.14	0.000	538020.26	1503291.76
	20	−282015.25	140986.43	0.959	−818737.77	254707.27
	21	2036667.06*	117321.65	0.000	1587615.80	2485718.32
	22	747866.71*	128158.54	0.000	259414.81	1236318.61
	23	−614603.48*	143402.85	0.011	−1160551.85	−68655.11
	24	1809717.01*	118347.91	0.000	1357009.56	2262424.47
	25	1286243.23*	125080.03	0.000	809144.39	1763342.07
	26	41258.90	138674.23	1.000	−486669.26	569187.06
	27	2257962.79*	116339.98	0.000	1812391.29	2703534.29
9	1	−1261452.03*	101091.25	0.000	−1647040.14	−875863.92
	2	−2596013.18*	119780.14	0.000	−3054953.38	−2137072.98
	3	−225957.23	88486.65	0.688	−562826.58	110912.12
	4	−1541422.65*	103203.59	0.000	−1935255.45	−1147589.85
	5	−2939065.31*	122869.81	0.000	−3410190.42	−2467940.20
	6	−457873.52*	89929.63	0.001	−800264.95	−115482.10
	7	−989121.61*	99058.41	0.000	−1366792.20	−611451.03
	8	−2262903.16*	116805.74	0.000	−2710123.45	−1815682.87
	10	−1266635.42*	102268.93	0.000	−1656818.10	−876452.74
	11	−2628988.97*	121899.09	0.000	−3096284.75	−2161693.19
	12	−214875.68	89123.49	0.785	−554179.86	124428.51
	13	−1553482.28*	104508.18	0.000	−1952414.91	−1154549.65
	14	−2982170.89*	125196.93	0.000	−3462479.53	−2501862.25
	15	−451496.30*	90643.14	0.001	−796624.75	−106367.85
	16	−987872.75*	100116.83	0.000	−1369663.49	−606082.02
	17	−2286393.73*	118728.86	0.000	−2741190.31	−1831597.15
	18	15477.21	87663.58	1.000	−318250.89	349205.31
	19	−1242247.15*	99831.93	0.000	−1622928.36	−861565.94
	20	−2544918.41*	117557.60	0.000	−2995100.11	−2094736.71
	21	−226236.10	87789.44	0.671	−560444.15	107971.95
	22	−1515036.45*	101819.58	0.000	−1903465.43	−1126607.47
	23	−2877506.64*	120444.98	0.000	−3339067.93	−2415945.35
	24	−453186.14*	89156.29	0.001	−792615.82	−113756.47
	25	−976659.93*	97916.44	0.000	−1349890.89	−603428.97
	26	−2221644.26*	114774.37	0.000	−2660867.50	−1782421.02
	27	−4940.37	86473.16	1.000	−334137.08	324256.34

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
10	1	5183.39	114421.05	1.000	−430410.41	440777.19
	2	−1329377.76*	131225.07	0.000	−1829679.99	−829075.53
	3	1040678.19*	103453.49	0.000	646149.29	1435207.09
	4	−274787.23	116291.52	0.814	−717500.08	167925.62
	5	−1672429.89*	134051.22	0.000	−2183739.16	−1161120.62
	6	808761.89*	104690.38	0.000	409672.84	1207850.95
	7	277513.80	112629.06	0.751	−151286.22	706313.83
	8	−996267.74*	128515.83	0.000	−1486044.17	−506491.31
	9	1266635.42*	102268.93	0.000	876452.74	1656818.10
	11	−1362353.55*	133162.03	0.000	−1870196.94	−854510.16
	12	1051759.74*	103998.72	0.000	655223.35	1448296.14
	13	−286846.86	117450.82	0.765	−733985.12	160291.40
	14	−1715535.47*	136187.40	0.000	−2235180.74	−1195890.20
	15	815139.12*	105303.92	0.000	413780.25	1216497.99
	16	278762.67	113561.07	0.757	−153567.37	711092.70
	17	−1019758.31*	130266.19	0.000	−1516332.03	−523184.59
	18	1282112.63*	102750.38	0.000	890166.05	1674059.21
	19	24388.27	113309.98	1.000	−406990.02	455766.56
	20	−1278282.99*	129199.56	0.000	−1770713.20	−785852.78
	21	1040399.32*	102857.78	0.000	648058.75	1432739.89
	22	−248401.03	115065.04	0.912	−686442.81	189640.75
	23	−1610871.22*	131832.20	0.000	−2113535.85	−1108206.59
	24	813449.27*	104026.83	0.000	416809.28	1210089.27
	25	289975.49	111626.01	0.656	−135034.10	714985.08
	26	−955008.84*	126672.39	0.000	−1437639.62	−472378.06
	27	1261695.05*	101736.64	0.000	873458.33	1649931.77
11	1	1367536.94*	132259.72	0.000	863006.12	1872067.76
	2	32975.79	147039.03	1.000	−526796.28	592747.86
	3	2403031.74*	122894.59	0.000	1932232.69	2873830.79
	4	1087566.32*	133881.19	0.000	577076.09	1598056.55
	5	−310076.34	149566.65	0.940	−879463.56	259310.88
	6	2171115.45*	123937.61	0.000	1696626.69	2645604.20
	7	1639867.36*	130712.51	0.000	1140993.90	2138740.81
	8	366085.81	144626.32	0.704	−184546.35	916717.97
	9	2628988.97*	121899.09	0.000	2161693.19	3096284.75
	10	1362353.55*	133162.03	0.000	854510.16	1870196.94
	12	2414113.30*	123353.92	0.000	1941691.71	2886534.88
	13	1075506.69*	134889.40	0.000	561296.18	1589717.20
	14	−353181.92	151484.20	0.832	−929888.35	223524.51
	15	2177492.67*	124456.30	0.000	1701162.05	2653823.29
	16	1641116.22*	131516.43	0.000	1139306.88	2142925.55
	17	342595.24	146183.91	0.825	−213932.97	899123.45
	18	2644466.18*	122303.29	0.000	2175750.18	3113182.18
	19	1386741.82*	131299.68	0.000	885724.84	1887758.80
	20	84070.56	145234.23	1.000	−468860.77	637001.89
	21	2402752.87*	122393.54	0.000	1933719.35	2871786.39
	22	1113952.52*	132817.24	0.000	607376.04	1620529.00
	23	−248517.67	147581.11	0.995	−810348.55	313313.21
	24	2175802.83*	123377.62	0.000	1703297.43	2648308.22
	25	1652329.04*	129849.23	0.000	1156599.11	2148058.97
	26	407344.71	142990.73	0.467	−137114.52	951803.94
	27	2624048.60*	121452.87	0.000	2158317.14	3089780.06

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
12	1	−1046576.35*	102840.85	0.000	−1438609.19	−654543.52
	2	−2381137.50*	121260.39	0.000	−2845313.84	−1916961.17
	3	−11081.55	90480.31	1.000	−355533.00	333369.89
	4	−1326546.97*	104917.97	0.000	−1726662.89	−926431.06
	5	−2724189.63*	124313.28	0.000	−3200391.48	−2247987.79
	6	−242997.85	91892.00	0.621	−592824.51	106828.81
	7	−774245.94*	100843.28	0.000	−1158523.96	−389967.92
	8	−2048027.48*	118323.21	0.000	−2500646.71	−1595408.26
	9	214875.68	89123.49	0.785	−124428.51	554179.86
	10	−1051759.74*	103998.72	0.000	−1448296.14	−655223.35
	11	−2414113.29*	123353.92	0.000	−2886534.88	−1941691.71
	13	−1338606.60*	106201.50	0.000	−1743726.17	−933487.04
	14	−2767295.21*	126613.86	0.000	−3252566.60	−2282023.83
	15	−236620.62	92590.38	0.687	−589113.40	115872.15
	16	−772997.08*	101883.16	0.000	−1161309.63	−384684.53
	17	−2071518.05*	120222.05	0.000	−2531607.23	−1611428.88
	18	230352.89	89675.54	0.677	−111043.20	571748.97
	19	−1027371.47*	101603.22	0.000	−1414597.36	−640145.59
	20	−2330042.73*	119065.48	0.000	−2785581.31	−1874504.16
	21	−11360.42	89798.59	1.000	−353223.22	330502.37
	22	−1300160.77*	103556.88	0.000	−1694977.92	−905343.63
	23	−2662630.96*	121917.15	0.000	−3129393.29	−2195868.64
	24	−238310.47	91135.31	0.643	−585253.51	108632.57
	25	−761784.25*	99721.76	0.000	−1141717.29	−381851.22
	26	−2006768.58*	116318.34	0.000	−2451507.59	−1562029.58
	27	209935.31	88512.18	0.809	−127056.37	546926.98
13	1	292030.25	116426.81	0.720	−151229.67	735290.17
	2	−1042530.90*	132977.61	0.000	−1549316.32	−535745.48
	3	1327525.05*	105667.63	0.000	924360.84	1730689.26
	4	12059.63	118265.56	1.000	−438170.51	462289.77
	5	−1385583.03*	135767.28	0.000	−1903207.38	−867958.68
	6	1095608.75*	106878.91	0.000	688002.06	1503215.45
	7	564360.67*	114666.17	0.001	127747.53	1000973.80
	8	−709420.88*	130304.81	0.000	−1205849.28	−212992.48
	9	1553482.28*	104508.18	0.000	1154549.65	1952414.91
	10	286846.86	117450.82	0.765	−160291.40	733985.12
	11	−1075506.69*	134889.40	0.000	−1589717.20	−561296.18
	12	1338606.60*	106201.50	0.000	933487.04	1743726.17
	14	−1428688.61*	137876.87	0.000	−1954527.02	−902850.20
	15	1101985.98*	107479.95	0.000	692167.06	1511804.90
	16	565609.53*	115581.75	0.001	125543.36	1005675.69
	17	−732911.45*	132031.45	0.000	−1236027.23	−229795.67
	18	1568959.49*	104979.35	0.000	1168309.79	1969609.19
	19	311235.13	115335.06	0.580	−127899.90	750370.16
	20	−991436.13*	130979.20	0.000	−1490475.00	−492397.26
	21	1327246.18*	105084.48	0.000	926212.89	1728279.47
	22	38445.83	117059.77	1.000	−407210.31	484101.97
	23	−1324024.36*	133576.77	0.000	−1833135.39	−814913.33
	24	1100296.13*	106229.03	0.000	695075.63	1505516.64
	25	576822.35*	113681.10	0.001	143915.46	1009729.24
	26	−668161.98*	128487.04	0.000	−1157564.06	−178759.90
	27	1548541.91*	103987.35	0.000	1151503.10	1945580.72

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
14	1	1720718.86*	135305.26	0.000	1204293.95	2237143.77
	2	386157.71	149784.37	0.670	−184114.52	956429.94
	3	2756213.66*	126166.40	0.000	2272513.50	3239913.82
	4	1440748.24*	136890.67	0.000	918528.97	1962967.51
	5	43105.58	152266.43	1.000	−536568.00	622779.16
	6	2524297.37*	127182.60	0.000	2037023.75	3011570.98
	7	1993049.28*	133793.28	0.000	1482121.53	2503977.02
	8	719267.73*	147416.60	0.001	157924.78	1280610.68
	9	2982170.89*	125196.93	0.000	2501862.25	3462479.53
	10	1715535.47*	136187.40	0.000	1195890.20	2235180.74
	11	353181.92	151484.20	0.832	−223524.51	929888.35
	12	2767295.22*	126613.86	0.000	2282023.83	3252566.60
	13	1428688.61*	137876.87	0.000	902850.20	1954527.02
	15	2530674.59*	127688.11	0.000	2041616.58	3019732.60
	16	1994298.14*	134578.80	0.000	1480518.08	2508078.19
	17	695777.16*	148945.01	0.003	128674.88	1262879.44
	18	2997648.10*	125590.51	0.000	2515964.71	3479331.49
	19	1739923.74*	134366.99	0.000	1226913.57	2252933.91
	20	437252.48	148013.05	0.391	−126335.96	1000840.92
	21	2755934.79*	125678.40	0.000	2273944.01	3237925.57
	22	1467134.44*	135850.29	0.000	948720.95	1985547.93
	23	104664.25	150316.55	1.000	−467620.35	676948.85
	24	2528984.75*	126636.95	0.000	2043632.18	3014337.31
	25	2005510.96*	132950.00	0.000	1497636.22	2513385.70
	26	760526.63*	145812.31	0.000	205210.10	1315843.16
	27	2977230.52*	124762.50	0.000	2498435.87	3456025.17
15	1	−809955.73*	104160.56	0.000	−1206878.10	−413033.36
	2	−2144516.88*	122381.64	0.000	−2612684.96	−1676348.80
	3	225539.07	91977.54	0.758	−124631.52	575709.66
	4	−1089926.35*	106211.88	0.000	−1494812.67	−685040.03
	5	−2487569.01*	125407.23	0.000	−2967643.09	−2007494.93
	6	−6377.22	93366.59	1.000	−361816.81	349062.36
	7	−537625.31*	102188.80	0.000	−926912.51	−148338.12
	8	−1811406.86*	119472.02	0.000	−2268139.63	−1354674.09
	9	451496.30*	90643.14	0.001	106367.85	796624.75
	10	−815139.12*	105303.92	0.000	−1216497.99	−413780.25
	11	−2177492.67*	124456.30	0.000	−2653823.29	−1701162.05
	12	236620.63	92590.38	0.687	−115872.15	589113.40
	13	−1101985.98*	107479.95	0.000	−1511804.90	−692167.06
	14	−2530674.59*	127688.11	0.000	−3019732.60	−2041616.58
	16	−536376.45*	103215.12	0.000	−929635.26	−143117.65
	17	−1834897.43*	121352.89	0.000	−2299020.62	−1370774.24
	18	466973.51*	91185.99	0.000	119796.29	814150.73
	19	−790750.85*	102938.80	0.000	−1182939.79	−398561.91
	20	−2093422.11*	120207.20	0.000	−2553042.79	−1633801.43
	21	225260.20	91307.00	0.749	−122374.19	572894.59
	22	−1063540.15*	104867.58	0.000	−1463205.17	−663875.13
	23	−2426010.34*	123032.41	0.000	−2896738.10	−1955282.58
	24	−1689.84	92621.95	1.000	−354302.36	350922.67
	25	−525163.63*	101082.20	0.000	−910175.28	−140151.98
	26	−1770147.96*	117486.76	0.000	−2219087.78	−1321208.14
	27	446555.93*	90042.15	0.001	103691.64	789420.22

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
16	1	−273579.28	112501.66	0.772	−701864.87	154706.32
	2	−1608140.43*	129554.87	0.000	−2102295.74	−1113985.11
	3	761915.52*	101326.55	0.000	375662.95	1148168.10
	4	−553549.90*	114403.51	0.001	−989101.82	−117997.97
	5	−1951192.56*	132416.67	0.000	−2456518.31	−1445866.80
	6	529999.23*	102589.09	0.000	139068.09	920930.37
	7	−1248.86	110678.59	1.000	−422594.87	420097.15
	8	−1275030.41*	126809.95	0.000	−1758495.46	−791565.35
	9	987872.75*	100116.83	0.000	606082.02	1369663.49
	10	−278762.67	113561.07	0.757	−711092.70	153567.37
	11	−1641116.22*	131516.43	0.000	−2142925.55	−1139306.88
	12	772997.08*	101883.16	0.000	384684.53	1161309.63
	13	−565609.53*	115581.75	0.001	−1005675.69	−125543.36
	14	−1994298.14*	134578.80	0.000	−2508078.19	−1480518.08
	15	536376.45*	103215.12	0.000	143117.65	929635.26
	17	−1298520.98*	128583.53	0.000	−1788890.47	−808151.48
	18	1003349.96*	100608.57	0.000	619748.09	1386951.84
	19	−254374.40	111371.44	0.857	−678354.62	169605.83
	20	−1557045.66*	127502.83	0.000	−2043206.74	−1070884.57
	21	761636.65*	100718.26	0.000	377630.29	1145643.02
	22	−527163.70*	113156.57	0.003	−957948.54	−96378.85
	23	−1889633.89*	130169.79	0.000	−2386187.38	−1393080.39
	24	534686.61*	101911.86	0.000	146267.76	923105.46
	25	11212.82	109657.70	1.000	−406259.09	428684.74
	26	−1233771.51*	124941.34	0.000	−1709974.42	−757568.59
	27	982932.38*	99573.03	0.000	603140.26	1362724.51
17	1	1024941.70*	129343.67	0.000	531773.43	1518109.97
	2	−309619.45	144421.70	0.918	−859420.80	240181.90
	3	2060436.50*	119750.71	0.000	1602022.29	2518850.71
	4	744971.08*	131001.24	0.000	245677.22	1244264.94
	5	−652671.58*	146994.34	0.006	−1212299.31	−93043.85
	6	1828520.21*	120820.88	0.000	1366297.53	2290742.88
	7	1297272.12*	127761.16	0.000	809922.87	1784621.36
	8	23490.57	141964.52	1.000	−516962.18	563943.32
	9	2286393.73*	118728.86	0.000	1831597.15	2741190.31
	10	1019758.31*	130266.19	0.000	523184.59	1516332.03
	11	−342595.24	146183.91	0.825	−899123.45	213932.97
	12	2071518.05*	120222.05	0.000	1611428.88	2531607.23
	13	732911.45*	132031.45	0.000	229795.67	1236027.23
	14	−695777.16*	148945.01	0.003	−1262879.44	−128674.88
	15	1834897.43*	121352.89	0.000	1370774.24	2299020.62
	16	1298520.98*	128583.53	0.000	808151.48	1788890.47
	18	2301870.94*	119143.81	0.000	1845607.60	2758134.28
	19	1044146.58*	128361.83	0.000	554592.10	1533701.06
	20	−258524.68	142583.78	0.987	−801329.84	284280.48
	21	2060157.63*	119236.45	0.000	1603566.40	2516748.86
	22	771357.28*	129913.71	0.000	276085.83	1266628.73
	23	−591112.91*	144973.57	0.022	−1143019.00	−39206.82
	24	1833207.59*	120246.37	0.000	1373031.90	2293383.27
	25	1309733.80*	126877.79	0.000	825619.57	1793848.03
	26	64749.47	140297.90	1.000	−469384.88	598883.82
	27	2281453.36*	118270.67	0.000	1828272.66	2734634.06

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
18	1	−1276929.24*	101578.27	0.000	−1664306.82	−889551.66
	2	−2611490.39*	120191.47	0.000	−3071881.59	−2151099.19
	3	−241434.44	89042.64	0.570	−580414.21	97545.33
	4	−1556899.86*	103680.70	0.000	−1952476.73	−1161322.99
	5	−2954542.52*	123270.83	0.000	−3427074.16	−2482010.88
	6	−473350.73*	90476.76	0.000	−817811.00	−128890.47
	7	−1004598.82*	99555.38	0.000	−1384104.64	−625093.01
	8	−2278380.37*	117227.50	0.000	−2727097.12	−1829663.62
	9	−15477.21	87663.58	1.000	−349205.31	318250.89
	10	−1282112.63*	102750.38	0.000	−1674059.21	−890166.05
	11	−2644466.18*	122303.29	0.000	−3113182.18	−2175750.18
	12	−230352.89	89675.54	0.677	−571748.97	111043.20
	13	−1568959.49*	104979.35	0.000	−1969609.19	−1168309.79
	14	−2997648.10*	125590.51	0.000	−3479331.49	−2515964.71
	15	−466973.51*	91185.99	0.000	−814150.73	−119796.29
	16	−1003349.96*	100608.57	0.000	−1386951.84	−619748.09
	17	−2301870.94*	119143.81	0.000	−2758134.28	−1845607.60
	19	−1257724.36*	100325.07	0.000	−1640223.15	−875225.57
	20	−2560395.62*	117976.67	0.000	−3012062.03	−2108729.21
	21	−241713.31	88349.83	0.552	−578052.38	94625.76
	22	−1530513.66*	102303.14	0.000	−1920716.22	−1140311.10
	23	−2892983.85*	120854.04	0.000	−3355986.35	−2429981.35
	24	−468663.35*	89708.14	0.000	−810183.99	−127142.72
	25	−992137.14*	98419.18	0.000	−1367229.98	−617044.30
	26	−2237121.47*	115203.57	0.000	−2677873.81	−1796369.13
	27	−20417.58	87042.02	1.000	−351785.08	310949.92
19	1	−19204.88	112248.20	1.000	−446527.60	408117.84
	2	−1353766.03*	129334.84	0.000	−1847113.97	−860418.09
	3	1016289.92*	101045.06	0.000	631131.15	1401448.69
	4	−299175.50	114154.28	0.639	−733784.36	135433.36
	5	−1696818.16*	132201.39	0.000	−2201358.30	−1192278.02
	6	784373.63*	102311.08	0.000	394520.20	1174227.05
	7	253125.54	110420.95	0.853	−167237.96	673489.03
	8	−1020656.01*	126585.14	0.000	−1503291.76	−538020.26
	9	1242247.15*	99831.93	0.000	861565.94	1622928.36
	10	−24388.27	113309.98	1.000	−455766.56	406990.02
	11	−1386741.82*	131299.68	0.000	−1887758.80	−885724.84
	12	1027371.47*	101603.22	0.000	640145.59	1414597.36
	13	−311235.13	115335.06	0.580	−750370.16	127899.90
	14	−1739923.74*	134366.99	0.000	−2252933.91	−1226913.57
	15	790750.85*	102938.80	0.000	398561.91	1182939.79
	16	254374.40	111371.44	0.857	−169605.83	678354.62
	17	−1044146.58*	128361.83	0.000	−1533701.06	−554592.10
	18	1257724.36*	100325.07	0.000	875225.57	1640223.15
	20	−1302671.26*	127279.24	0.000	−1788008.68	−817333.84
	21	1016011.05*	100435.07	0.000	633106.35	1398915.75
	22	−272789.30	112904.58	0.782	−702618.17	157039.57
	23	−1635259.49*	129950.80	0.000	−2131010.38	−1139508.60
	24	789061.00*	101631.99	0.000	401728.46	1176393.55
	25	265587.22	109397.65	0.774	−150890.90	682065.34
	26	−979397.11*	124713.17	0.000	−1454755.17	−504039.05
	27	1237306.78*	99286.58	0.000	858631.34	1615982.22

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
20	1	1283466.38*	128269.37	0.000	794477.03	1772455.73
	2	−51094.77	143460.36	1.000	−597244.51	495054.97
	3	2318961.18*	118589.54	0.000	1865117.77	2772804.59
	4	1003495.76*	129940.65	0.000	508317.50	1498674.02
	5	−394146.90	146049.93	0.580	−950202.46	161908.66
	6	2087044.88*	119670.09	0.000	1629347.28	2544742.49
	7	1555796.80*	126673.44	0.000	1072688.17	2038905.42
	8	282015.25	140986.43	0.959	−254707.27	818737.77
	9	2544918.41*	117557.60	0.000	2094736.71	2995100.11
	10	1278282.99*	129199.56	0.000	785852.78	1770713.20
	11	−84070.56	145234.23	1.000	−637001.89	468860.77
	12	2330042.73*	119065.48	0.000	1874504.16	2785581.31
	13	991436.13*	130979.20	0.000	492397.26	1490475.00
	14	−437252.48	148013.05	0.391	−1000840.92	126335.96
	15	2093422.11*	120207.20	0.000	1633801.43	2553042.79
	16	1557045.66*	127502.83	0.000	1070884.57	2043206.74
	17	258524.68	142583.78	0.987	−284280.48	801329.84
	18	2560395.62*	117976.67	0.000	2108729.21	3012062.03
	19	1302671.26*	127279.24	0.000	817333.84	1788008.68
	21	2318682.31*	118070.22	0.000	1866684.01	2770680.61
	22	1029881.96*	128844.16	0.000	538767.50	1520996.42
	23	−332588.23	144015.92	0.844	−880860.00	215683.54
	24	2091732.26*	119090.03	0.000	1636106.14	2547358.39
	25	1568258.48*	125782.43	0.000	1088419.84	2048097.12
	26	323274.15	139308.10	0.838	−207075.07	853623.37
	27	2539978.04*	117094.83	0.000	2091432.11	2988523.97
21	1	−1035215.93*	101686.92	0.000	−1422993.18	−647438.68
	2	−2369777.08*	120283.30	0.000	−2830492.66	−1909061.50
	3	278.87	89166.56	1.000	−339171.71	339729.45
	4	−1315186.55*	103787.14	0.000	−1711153.01	−919220.09
	5	−2712829.21*	123360.37	0.000	−3185675.32	−2239983.10
	6	−231637.42	90598.72	0.686	−576559.29	113284.44
	7	−762885.51*	99666.23	0.000	−1142801.16	−382969.87
	8	−2036667.06*	117321.65	0.000	−2485718.32	−1587615.80
	9	226236.10	87789.44	0.671	−107971.95	560444.15
	10	−1040399.32*	102857.78	0.000	−1432739.89	−648058.75
	11	−2402752.87*	122393.54	0.000	−2871786.39	−1933719.35
	12	11360.42	89798.59	1.000	−330502.37	353223.22
	13	−1327246.18*	105084.48	0.000	−1728279.47	−926212.89
	14	−2755934.79*	125678.40	0.000	−3237925.57	−2273944.01
	15	−225260.20	91307.00	0.749	−572894.59	122374.19
	16	−761636.65*	100718.26	0.000	−1145643.02	−377630.29
	17	−2060157.63*	119236.45	0.000	−2516748.86	−1603566.40
	18	241713.31	88349.83	0.552	−94625.76	578052.38
	19	−1016011.05*	100435.07	0.000	−1398915.75	−633106.35
	20	−2318682.31*	118070.22	0.000	−2770680.61	−1866684.01
	22	−1288800.35*	102411.02	0.000	−1679399.05	−898201.65
	23	−2651270.54*	120945.37	0.000	−3114595.23	−2187945.85
	24	−226950.04	89831.14	0.708	−568937.18	115037.09
	25	−750423.83*	98531.31	0.000	−1125932.41	−374915.25
	26	−1995408.16*	115299.37	0.000	−2436502.28	−1554314.04
	27	221295.73	87168.78	0.699	−110555.93	553147.39

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
22	1	253584.42	114019.61	0.885	−180478.50	687647.34
	2	−1080976.73*	130875.18	0.000	−1579988.60	−581964.86
	3	1289079.22*	103009.31	0.000	896280.32	1681878.12
	4	−26386.20	115896.56	1.000	−467598.58	414826.18
	5	−1424028.86	133708.72	0.000	−1934081.74	−913975.98
	6	1057162.92*	104251.47	0.000	659779.59	1454546.26
	7	525914.84*	112221.20	0.003	98675.77	953153.90
	8	−747866.71*	128158.54	0.000	−1236318.61	−259414.81
	9	1515036.45*	101819.58	0.000	1126607.47	1903465.43
	10	248401.03	115065.04	0.912	−189640.75	686442.81
	11	−1113952.52*	132817.24	0.000	−1620529.00	−607376.04
	12	1300160.77*	103556.88	0.000	905343.63	1694977.92
	13	−38445.83	117059.77	1.000	−484101.97	407210.31
	14	−1467134.44*	135850.29	0.000	−1985547.93	−948720.95
	15	1063540.15*	104867.58	0.000	663875.13	1463205.17
	16	527163.70*	113156.57	0.003	96378.85	957948.54
	17	−771357.28*	129913.71	0.000	−1266628.73	−276085.83
	18	1530513.66*	102303.14	0.000	1140311.10	1920716.22
	19	272789.30	112904.58	0.782	−157039.57	702618.17
	20	−1029881.96*	128844.16	0.000	−1520996.42	−538767.50
	21	1288800.35*	102411.02	0.000	898201.65	1679399.05
	23	−1362470.19*	131483.93	0.000	−1863851.89	−861088.49
	24	1061850.30*	103585.11	0.000	666929.00	1456771.61
	25	538376.52*	111214.48	0.001	114945.13	961807.91
	26	−706607.81*	126309.88	0.000	−1187889.89	−225325.73
	27	1510096.08*	101284.93	0.000	1123623.84	1896568.32
23	1	1616054.61*	130920.72	0.000	1116744.62	2115364.60
	2	281493.46	145835.79	0.972	−273689.80	836676.72
	3	2651549.41*	121452.39	0.000	2186433.28	3116665.54
	4	1336083.99*	132558.58	0.000	830739.31	1841428.67
	5	−61558.67	148383.92	1.000	−626453.27	503335.93
	6	2419633.12*	122507.69	0.000	1950773.69	2888492.54
	7	1888385.03*	129357.50	0.000	1394805.60	2381964.45
	8	614603.48*	143402.85	0.011	68655.11	1160551.85
	9	2877506.64*	120444.98	0.000	2415945.35	3339067.93
	10	1610871.22*	131832.20	0.000	1108206.59	2113535.85
	11	248517.67	147581.11	0.995	−313313.21	810348.55
	12	2662630.97*	121917.15	0.000	2195868.64	3129393.29
	13	1324024.36*	133576.77	0.000	814913.33	1833135.39
	14	−104664.25	150316.55	1.000	−676948.85	467620.35
	15	2426010.34*	123032.41	0.000	1955282.58	2896738.10
	16	1889633.89*	130169.79	0.000	1393080.39	2386187.38
	17	591112.91*	144973.57	0.022	39206.82	1143019.00
	18	2892983.85*	120854.04	0.000	2429981.35	3355986.35
	19	1635259.49*	129950.80	0.000	1139508.60	2131010.38
	20	332588.23	144015.92	0.844	−215683.54	880860.00
	21	2651270.54*	120945.37	0.000	2187945.85	3114595.23
	22	1362470.19*	131483.93	0.000	861088.49	1863851.89
	24	2424320.50*	121941.13	0.000	1957473.14	2891167.85
	25	1900846.71*	128485.12	0.000	1410452.21	2391241.21
	26	655862.38*	141753.14	0.003	116153.09	1195571.67
	27	2872566.27*	119993.34	0.000	2412592.55	3332539.99

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
24	1	−808265.88*	102869.27	0.000	−1200403.79	−416127.98
	2	−2142827.03*	121284.50	0.000	−2607088.98	−1678565.09
	3	227228.92	90512.62	0.719	−117345.72	571803.55
	4	−1088236.50*	104945.84	0.000	−1488454.90	−688018.11
	5	−2485879.16*	124336.79	0.000	−2962164.03	−2009594.30
	6	−4687.38	91923.81	1.000	−354634.90	345260.14
	7	−535935.47*	100872.27	0.000	−920321.16	−151549.78
	8	−1809717.01*	118347.91	0.000	−2262424.47	−1357009.56
	9	453186.15*	89156.29	0.001	113756.47	792615.82
	10	−813449.27*	104026.83	0.000	−1210089.27	−416809.28
	11	−2175802.83*	123377.62	0.000	−2648308.22	−1703297.43
	12	238310.47	91135.31	0.643	−108632.57	585253.51
	13	−1100296.13*	106229.03	0.000	−1505516.64	−695075.63
	14	−2528984.74*	126636.95	0.000	−3014337.31	−2043632.18
	15	1689.84	92621.95	1.000	−350922.67	354302.36
	16	−534686.61*	101911.86	0.000	−923105.46	−146267.76
	17	−1833207.58*	120246.37	0.000	−2293383.27	−1373031.90
	18	468663.36*	89708.14	0.000	127142.72	810183.99
	19	−789061.00*	101631.99	0.000	−1176393.55	−401728.46
	20	−2091732.26*	119090.03	0.000	−2547358.39	−1636106.14
	21	226950.04	89831.14	0.708	−115037.09	568937.18
	22	−1061850.30*	103585.11	0.000	−1456771.61	−666929.00
	23	−2424320.49*	121941.13	0.000	−2891167.85	−1957473.14
	25	−523473.79*	99751.07	0.000	−903516.01	−143431.56
	26	−1768458.12*	116343.47	0.000	−2213287.24	−1323628.99
	27	448245.78*	88545.21	0.001	111127.54	785364.01
25	1	−284792.10	110548.05	0.672	−705670.28	136086.08
	2	−1619353.25*	127862.10	0.000	−2107312.57	−1131393.93
	3	750702.70*	99153.01	0.000	372885.86	1128519.54
	4	−564762.72*	112482.95	0.001	−993062.13	−136463.31
	5	−1962405.38*	130760.94	0.000	−2461703.90	−1463106.86
	6	518786.41*	100442.88	0.000	136164.25	901408.56
	7	−12461.68	108692.23	1.000	−426246.54	401323.17
	8	−1286243.23*	125080.03	0.000	−1763342.07	−809144.39
	9	976659.93*	97916.44	0.000	603428.97	1349890.89
	10	−289975.49	111626.01	0.656	−714985.08	135034.10
	11	−1652329.04*	129849.23	0.000	−2148058.97	−1156599.11
	12	761784.25*	99721.76	0.000	381851.22	1141717.29
	13	−576822.35*	113681.10	0.001	−1009729.24	−143915.46
	14	−2005510.96*	132950.00	0.000	−2513385.70	−1497636.22
	15	525163.63*	101082.20	0.000	140151.98	910175.28
	16	−11212.82	109657.70	1.000	−428684.74	406259.09
	17	−1309733.80*	126877.79	0.000	−1793848.03	−825619.57
	18	992137.14*	98419.18	0.000	617044.30	1367229.98
	19	−265587.22	109397.65	0.774	−682065.34	150890.90
	20	−1568258.48*	125782.43	0.000	−2048097.12	−1088419.84
	21	750423.83*	98531.31	0.000	374915.25	1125932.41
	22	−538376.52*	111214.48	0.001	−961807.91	−114945.13
	23	−1900846.71*	128485.12	0.000	−2391241.21	−1410452.21
	24	523473.79*	99751.07	0.000	143431.56	903516.01
	26	−1244984.33*	123185.18	0.000	−1714700.19	−775268.47
	27	971719.56*	97360.36	0.000	600543.78	1342895.34

TABLE 8: Continued.

(I) experiment	(J) experiment	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
26	1	960192.23*	125723.50	0.000	481088.89	1439295.57
	2	−374368.92	141188.68	0.616	−911914.42	163176.58
	3	1995687.03*	115831.11	0.000	1552693.09	2438680.97
	4	680221.61*	127428.18	0.000	194774.57	1165668.65
	5	−717421.05*	143819.16	0.001	−1265064.13	−169777.97
	6	1763770.74*	116937.16	0.000	1316809.66	2210731.81
	7	1232522.65*	124094.83	0.000	759451.09	1705594.20
	8	−41258.90	138674.23	1.000	−569187.06	486669.26
	9	2221644.26*	114774.37	0.000	1782421.02	2660867.50
	10	955008.84*	126672.39	0.000	472378.06	1437639.62
	11	−407344.71	142990.73	0.467	−951803.94	137114.52
	12	2006768.59*	116318.34	0.000	1562029.58	2451507.59
	13	668161.98*	128487.04	0.000	178759.90	1157564.06
	14	−760526.63*	145812.31	0.000	−1315843.16	−205210.10
	15	1770147.96*	117486.76	0.000	1321208.14	2219087.78
	16	1233771.51*	124941.34	0.000	757568.59	1709974.42
	17	−64749.47	140297.90	1.000	−598883.82	469384.88
	18	2237121.47*	115203.57	0.000	1796369.13	2677873.81
	19	979397.11*	124713.17	0.000	504039.05	1454755.17
	20	−323274.15	139308.10	0.838	−853623.37	207075.07
	21	1995408.16*	115299.37	0.000	1554314.04	2436502.28
	22	706607.81*	126309.88	0.000	225325.73	1187889.89
	23	−655862.38*	141753.14	0.003	−1195571.67	−116153.09
	24	1768458.12*	116343.47	0.000	1323628.99	2213287.24
	25	1244984.33*	123185.18	0.000	775268.47	1714700.19
	27	2216703.89*	114300.33	0.000	1779165.65	2654242.13
27	1	−1256511.66*	100552.72	0.000	−1640125.31	−872898.01
	2	−2591072.81*	119325.99	0.000	−3048414.59	−2133731.03
	3	−221016.86	87870.91	0.716	−555552.92	113519.20
	4	−1536482.28*	102676.15	0.000	−1928391.20	−1144573.36
	5	−2934124.94*	122427.13	0.000	−3403700.88	−2464549.00
	6	−452933.16*	89323.84	0.001	−793037.93	−112828.38
	7	−984181.25*	98508.77	0.000	−1359826.36	−608536.13
	8	−2257962.79*	116339.98	0.000	−2703534.29	−1812391.29
	9	4940.37	86473.16	1.000	−324256.34	334137.08
	10	−1261695.05*	101736.64	0.000	−1649931.77	−873458.33
	11	−2624048.60*	121452.87	0.000	−3089780.06	−2158317.14
	12	−209935.31	88512.18	0.809	−546926.98	127056.37
	13	−1548541.91*	103987.35	0.000	−1945580.72	−1151503.10
	14	−2977230.52*	124762.50	0.000	−3456025.17	−2498435.87
	15	−446555.93*	90042.15	0.001	−789420.22	−103691.64
	16	−982932.38*	99573.03	0.000	−1362724.51	−603140.26
	17	−2281453.36*	118270.67	0.000	−2734634.06	−1828272.66
	18	20417.58	87042.02	1.000	−310949.92	351785.08
	19	−1237306.78*	99286.58	0.000	−1615982.22	−858631.34
	20	−2539978.04*	117094.83	0.000	−2988523.97	−2091432.11
	21	−221295.73	87168.78	0.699	−553147.39	110555.93
	22	−1510096.08*	101284.93	0.000	−1896568.32	−1123623.84
	23	−2872566.27*	119993.34	0.000	−3332539.99	−2412592.55
	24	−448245.78*	88545.21	0.001	−785364.01	−111127.54
	25	−971719.56*	97360.36	0.000	−1342895.34	−600543.78
	26	−2216703.89*	114300.33	0.000	−2654242.13	−1779165.65

*The mean difference is significant at the 0.05 level.

TABLE 9: Differences.

| | Policy combination | | | | | | | | | | | | | | | | | | | | | | | | | |
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Dif1	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	×	✓	✓
Dif2	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓
Dif3	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓	×	✓	✓

Note. Dif1 refers to whether there is a significant difference in the economic effect with policy combination 1, Dif2 refers to whether there is a significant difference in the STI effect with policy Combination 1, and Dif3 refers to whether there is a significant difference in the economic effect and STI effect with policy Combination 1 at the same time.

5. Conclusions

5.1. Research Summary. STI has become an important driving force for national and regional economic and social development. As a tool for the government to guide and promote the promotion of scientific and technological competitiveness, STI policy has a significant impact on STI. Most of the existing studies focus on the effect of a single policy, and the studies on the effect of policy combination and its differences need to be further enriched and improved. Under this background, this study proposes a method combining system simulation experiment and ANOVA to study the differences of the combination effects of STI policies. The research is summarized as follows:

(1) On the whole, there are significant differences in both the economic effect of STI policy combination and the STI effect of STI policy combination.

(2) Specific to different combinations of STI policies, not all combinations of STI policies have significant differences in policy effects.

(3) There are significant differences in the economic effect of some STI policy combinations, but the differences in the STI effect are not significant. There are significant differences in STI effect among some STI policy combinations, but the differences in economic effect are not significant.

(4) This study not only points out whether there are significant differences in economic effect or STI effect among STI policy combinations, but also points out whether there are significant differences in economic effect and STI effect among STI policy combinations at the same time.

5.2. Policy Implications. This study gives us the following enlightenment:

Firstly, the empirical results show that there are significant differences in policy effects among some STI policy combinations, while there are no significant differences in policy effects among some STI policy combinations. Therefore, policymakers should fully consider the differences in policy effects among policy combinations when formulating policies, so as to avoid making useless efforts in formulating policies. In other words, policymakers should avoid formulating policies that have no significant difference in policy effects from the original policies.

Secondly, the empirical results show that the economic effect and STI effect of STI policy combinations are not always synchronous, that is, sometimes there are significant differences in one effect but not in the other. Therefore, when making policies, policy makers sometimes have to make a trade-off among a variety of policy effects.

5.3. Contribution and Prospect. The contributions of this study are as follows:

(1) The existing research mainly focuses on the effect of a single policy, and the research on the combination

effect of STI policy, especially the research on the difference of policy combination effect, needs to be further enriched and improved. This research enriches and expands the relevant literature research, makes up for the shortcomings of previous research, and provides an effective supplement for the existing research on STI policy.

(2) This study proposes a method combining system simulation experiment and ANOVA to study the differences of the combination effects of STI policies. This method based on the combination of system simulation and data analysis and mining enhances the modeling and solving ability of practical problems, provides a reference framework for similar research, and reflects the characteristics of this study in research methods.

This study provides enlightenment for policy makers to fully understand the effect of STI policy and make appropriate adjustment. At the same time, it enriches and expands the relevant literature research, which has certain theoretical and practical significance. However, this study also has some limitations. For example, only six kinds of STI policies are considered, and other STI policies are not deeply studied. In future research, more STI policies can be included in the research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Y. Zhang, Z. Geng, and Y. Wang, "The complex adaptation of regional science and technology innovation policy," *Chinese Journal of Systems Science*, vol. 23, no. 4, pp. 48–51, 2015.

[2] J. Peng, W. Zhong, and W. Sun, "Policy measurement, policy collaborative evolution and economic performance: an empirical study based on innovation policy," *Management World*, vol. 9, pp. 25–36, 2008.

[3] S. Surianto, A. Syahirul, R. D. Nindrea, and L. Trisnantoro, "Regional policy for disaster risk management in developing countries within the sendai framework: a systematic review," *Open Access Macedonian Journal of Medical Sciences*, vol. 7, no. 13, pp. 2213–2219, 2019.

[4] D. S. Krivosheev, E. F. Troitskiy, and E. F. Troitskiy, "The EU regional policy in wallonia (2000-17)," *Vestnik Tomskogo gosudarstvennogo universiteta*, vol. 444, no. 444, pp. 136–142, 2019.

[5] J. Shi, X. Chen, and Y. Liu, "Research of district strategy's promotion mechanism on Guangdong industry transfer," *Research of Finance and Accounting*, vol. 12, pp. 78–80, 2011.

[6] S. Yu, "The influence of regional economic policy on regional economic development," *China Business and Trade*, vol. 28, pp. 165-166, 2014.

[7] X. Lan, S. Wang, and Z. Zhu, "Analysis of the impact of regional economic policies on regional economic development," *Modern Business*, vol. 12, pp. 67-68, 2017.

[8] X. Lin, "The impact of EU's regional policy on the Portuguese national capability in terms of science, technology and innovation," *Forum on Science and Technology in China*, vol. 9, pp. 148–154, 2016.

[9] J. Wang and Y. Zhang, "A study on the influence mechanism of science, technology and innovation policy on regional innovation capacity based on system dynamics," *Science and Technology Management Research*, vol. 38, no. 8, pp. 52–57, 2018.

[10] A. Avdiushchenko and P. Zając, "Circular economy indicators as a supporting tool for European regional development policies," *Sustainability*, vol. 11, no. 11, p. 3025, 2019.

[11] M. O'Brien and S. Burrows, "Assessing the effectiveness of regional policy responses to mass redundancies: the case of the illawarra region, Australia," *Economic Papers: A journal of applied economics and policy*, vol. 38, no. 2, pp. 144–155, 2019.

[12] F. Gülal and A. Ayaita, "The impact of minimum wages on well-being: evidence from a quasi-experiment in Germany," *Journal of Happiness Studies*, vol. 21, no. 7, pp. 2669–2692, 2019.

[13] X Cheng, X. Bao, R. Zhang, and X. Shen, "Analysis on the efficiency of science and technology financial policy in beijing, tianjin and hebei region based on policy quantification," *Science and Technology Management Research*, vol. 38, no. 20, pp. 62–68, 2018.

[14] J. Deng and R. Long, "Empirical research on venture capital to promote technological innovation in China," *Technoeconomics & Management Research*, vol. 6, pp. 49–52, 2013.

[15] Yu. Research, "On the vertical gap of bank loans, government and enterprise science and technology investment contribution——estimation based on quantile regression of panel data," *Science & Technology Progress and Policy*, vol. 30, no. 4, pp. 28–32, 2013.

[16] Z. Fan, "The technological innovation effect, existing problems and strategic choice of intellectual property protection," *Scientific Management Research*, vol. 36, no. 5, pp. 27–30, 2018.

[17] Z. Wang and S. Deng, "Research on policy performance of coordinated development of intellectual property protection and science and technology innovation in Jiangxi Province," *Enterprise Economy*, vol. 36, no. 6, pp. 120–126, 2017.

[18] Y. Zhuang, Y. Li, and Y. Qin, "Innovation environment incentive and enterprise growth: an empirical study based on innovative enterprises in Central China," vol. 29, no. 1, pp. 86–90, 2012.

[19] I. C. Álvarez-Ayuso, C. Kao, and D. Romero-Jordán, "Long run effect of public grants and tax credits on R&D investment: a non-stationary panel data approach," *Economic Modelling*, vol. 75, pp. 93–104, 2018.

[20] B. Raffaello and P. Paolo, "The impact of R&D subsidies on firm innovation," *Research Policy*, vol. 45, no. 2, pp. 442–457, 2016.

[21] A. Mukherjee, M. Singh, and A. Žaldokas, "Do corporate taxes hinder innovation?" *Journal of Financial Economics*, vol. 124, no. 1, pp. 195–221, 2017.

[22] Å Cappelen, A Raknerud, and M Rybalka, "The effects of R&D tax credits on patenting and innovations," *Research Policy*, vol. 41, no. 2, pp. 334–345, 2012.

[23] M. Guerzoni and E. Raiteri, "Demand-side vs. supply-side technology policies: hidden treatment and new empirical evidence on the policy mix," *Research Policy*, vol. 44, no. 3, pp. 726–747, 2015.

[24] B. Aschhoff and W. Sofka, "Innovation on demand-Can public procurement drive market success of innovations?" *Research Policy*, vol. 38, no. 8, pp. 1235–1247, 2009.

[25] H. Song and S. Zhang, "China's government procurement policy system construction and development for innovation product," *Studies in Science of Science*, vol. 32, no. 11, pp. 1639–1645, 2014.

[26] K. Hu, H. Cai, and Q. Wu, "Has government procurement in China promoted technological innovation?" *Journal of Finance and Economics*, vol. 39, no. 9, pp. 134–144, 2013.

[27] I. Kalcheva, P. Mclemore, and S. Pant, "Innovation: the interplay between demand-side shock and supply-side environment," *Research Policy*, vol. 47, no. 2, pp. 440–461, 2018.

[28] S. Dou, J. Liu, and S. Pang, "Synergy effects of policy mix in technological innovation——based on perspective of supply side-demand side," *Science & Technology Progress and Policy*, vol. 36, no. 22, pp. 118–126, 2019.

[29] Y. Guo, J. Ge, C. Cheng, and S. Duan, "Supply mode of technology policies mix based on QCA method," *Soft Science*, vol. 33, no. 1, pp. 45–49, 2019.

[30] X. Qin, *Technology Subsidies, Tax Preferences and enterprise Innovation: The Effect of R&D Policies of Listed Companies*, Dongbei University of Finance and Economics, Dalian, China, 2018.

[31] P. M. Romer, "Increasing returns and long-run growth," *Journal of Political Economy*, vol. 94, no. 5, pp. 1002–1037, 1986.

[32] J. Fox, *Applied Regression Analysis and Generalized Linear Models*, Sage Publications Incorporated, California, 2016.

[33] D. G. Kleinbaum, L. L. Kupper, A. Nizam, and E. S. Rosenberg, *Applied Regression Analysis and Other Multivariable Methods*, Thomson Learning Incorporated, California, 2008.

[34] J. Zhang, "The method for multiple comparisons," *Journal of Evidence-Based Medicine*, vol. 8, no. 3, pp. 167–171, 2008.

[35] M. C. Shingala and D. A. Rajyaguru, "Comparison of post hoc tests for unequal variance," *International Journal of New Technologies in Science and Engineering*, vol. 2, no. 5, pp. 22–33, 2015.

[36] P. A. Games and J. F. Howell, "Pairwise multiple comparison procedures with unequal N's and/or variances: a Monte Carlo study," *Journal of Educational Statistics*, vol. 1, pp. 13–125, 1976.

[37] R. Rothwell and W. Zegveld, *Industrial Innovation and Public Policy: Preparing for the 1980s and the 1990s*, Greenwood Publishing Group, London, 1981.

Hindawi

Research Article

A Terrain Elevation Map Generation Method Based on Self-Attention Mechanism and Multifeature Sketch

Xingquan Cai ⓘ, Mengyao Xi ⓘ, Nu Yu ⓘ, Zhe Yang ⓘ, and Haiyan Sun ⓘ

School of Information Science and Technology, North China University of Technology, Beijing 100144, China

Correspondence should be addressed to Xingquan Cai; caixingquan@ncut.edu.cn

To address the issues of low efficiency in manual terrain feature map annotating and poor realism in terrain elevation map generation, this paper proposes a terrain elevation map generation method based on self-attention mechanism and multifeature sketch. Firstly, the proposed method extracts features from a terrain elevation map using an adaptive feature enhancement method. Afterwards, our method adds a self-attention mechanism to the generator and discriminator of conditional generative adversarial network to capture the global spatial features and generates a realistic terrain elevation map. Finally, a level of detail method is used to visualize the three-dimensional terrain, and an interactive terrain editing tool for roaming interaction is implemented. Experimental data show that the proposed method performs well in subjective visual performance and objective criteria and has obvious advantages over other current typical methods.

1. Introduction

Terrain is one of the most common elements in nature. Realistic 3D terrains have been widely used in 3D video games, military simulations, film industry, and other fields. For instance, developers in 3D video games can create a game world that brings immersive user experience to players; realistic battlefields can be simulated for virtual military training in military simulations; visual effect artists can create scenes with realistic terrains that bring amazing visual experience to the audience in the film industry. Hence, digital terrains with realistic features can play an important role in enhancing the realism of virtual scenes and improving user immersion.

Interactive terrain editing tools aim at generating realistic terrains efficiently based on the sketch input from users. These editing tools should provide rich functionalities that can support users with a little technical background. In terms of realism, generated terrains should have natural features of the terrain in the real world, where no traces of manual editing can be found. Since interactive terrain editing tools can be widely used in various applications, it is important to investigate how to efficiently generate terrain elevation maps with realistic natural features.

In recent years, many researchers have investigated terrain editing. Existing terrain editing methods can be classified into three categories: terrain editing based on fractal noises, terrain editing based on physical erosion simulations, and terrain editing based on real samples. Specifically, for terrain editing methods based on fractal noises, a terrain is simulated by generating fractal noises based on user input parameters. The process is hardly controllable, which is difficult for users to get results that meet their demands. As for the methods based on physical erosion simulations, users need to edit a terrain with the knowledge of physical erosion theory, which is difficult for nonprofessional users. Finally, for the methods based on real samples, despite the fact that generated terrains can retain some natural features of real terrain, there is still much room for optimization in terms of realism and interaction experience. In order to enable users to perform interactive editing flexibly and generate realistic terrain, this paper draws on the methods based on real samples, which not only ensures that the generated results have the features of real samples but also enhances the realism of the generated results further. At the same time, the proposed method can improve computing efficiency, meet real-time interaction requirements,

and enhance the user's interactive experience by visualizing the three-dimensional terrain. Therefore, it is necessary to investigate interactive terrain editing techniques based on multifeature sketches.

The contributions are summarized as follows:

(1) We propose a terrain feature extraction method based on adaptive feature enhancement to solve the problem of low efficiency of manual annotation.

(2) We propose a terrain elevation map generation method based on self-attention mechanism. The conditional generative adversarial network is used to generate a realistic terrain elevation map.

(3) We propose a three-dimensional terrain visualization method based on Levels of Detail, which improves the fluency of user interaction.

(4) We realize an interactive terrain editing tool based on multifeature sketches. The tool runs stably and has a good user interaction experience.

2. Related Work

Interactive terrain editing tools can generate terrain elevation maps based on various terrain features from user input. Relevant scholars proposed a method for interactive generation and display of 3D terrain [1]. The method can automatically generate terrain that meets the user's needs, but the feedback of the interactive operation is yet to be studied. The researchers gradually generated the terrain by expanding the details of the user sketch [2]. The algorithm is computationally efficient and highly interactive. Related scholars used terrain elevation entropy, extracted features from Digital Elevation Model (DEM) data to generate terrain frames, and accelerated mapping using GPU to improve the efficiency of terrain generation during the editing of terrain [3]. Given the high efficiency of the GPU, a terrain editing algorithm based on CUDA architecture was proposed [4], which is applicable to large-scale terrain scene creation. With the development of deep learning, the researchers applied this to the terrain editing problem and proposed a real-time interactive terrain editing method [5]. The results obtained by this method conform to the feature distribution of real terrain samples with high fidelity. Related scholars implemented interactive terrain editing by providing the user with functions related to plate tectonics as well as erosion simulation [6]. The results of terrain editing obtained by this method are in accordance with the theories related to geosciences and are visually closer to the real terrain. Researchers have proved that the approach with machine learning could enhance the realism of procedurally generated terrain [7]. The method used Deep Convolutional Generative Adversarial Network (DCGAN) to generate height maps, and the generated results outperformed the noise-generated elevation map results, achieving a more realistic terrain elevation map editing. The authors proposed a generation method based on implicit feature representation of terrain [8]. This method can create large-scale terrain with less memory consumption and can implement special

terrain. To solve the problem of less efficient manual editing, a semiautomatic procedural terrain generation method for terrain editing tasks has been proven to work [9].

Specifically, terrain feature extraction method, terrain elevation map generation method, and terrain visualization method will be introduced as follows.

For terrain feature extractions, the primary goal is to extract ridge lines and valley lines efficiently and accurately. A Profile recognition and Polygon breaking Algorithm (PPA) was firstly proposed in [10]. This method can efficiently extract terrain ridge lines and valley lines that were consistent with the actual observation. The researchers proposed the necessity of effectively selecting terrain elevation profiles based on DEM data, which can improve the accuracy of traditional elevation profile extraction methods [11]. Based on the research [10], relevant scholars proved that morphological correlation algorithm could be used to complete the region refinement after polygon construction [12]. Related researchers proposed a terrain feature extraction algorithm to control feature significance, which can obtain extraction results of terrain feature lines that were in line with human eye observations [13]. Inspired by PPA, we optimize the process of removing redundant branches and leaves in a feature-connected graph when extracting ridge lines and valley lines. We propose a terrain feature extraction method based on adaptive feature enhancement, which can extract ridge lines, valley lines, and peak points in large-scale data set production tasks.

For terrain elevation map generations, deep learning algorithms become mainstream in the field of image generation. Generative Adversarial Network was proposed [14], which can create a model for generating high-definition images. Subsequently, the researchers proposed Conditional Generative Adversarial Network that added constraints to the Generative Adversarial Network, which can use category labels as a condition to get generation results [15]. The scholars proposed Deep Convolutional Generative Adversarial Network (DCGAN) [16], which introduced convolutional operations into the structure of Generative Adversarial Network for the first time. This method can capture the features of input data well and realize supervised learning. Based on the research in [15], related scholars proposed a method where an input image can be mapped to an output image using a Conditional Generative Adversarial Network, which can effectively convert picture labels to pictures [17]. The researchers added a self-attention mechanism to the network's construction, which can capture structural global features more easily [18]. Inspired by the work [17, 18], we apply a Conditional Generative Adversarial Network in terrain elevation map generation tasks. In addition, a self-attention mechanism is used to capture the global structure of terrain feature map, which can improve the performance of generator and discriminator to generate realistic terrains.

For terrain visualizations, terrain mesh generation and texture mapping are two key issues. For terrain mesh generation, it is mainly based on Levels of Detail (LOD) techniques. Related scholars proposed the concept of LOD for the first time and proposed a method for reducing the

complexity of model mesh in a recursive way [19]. The researchers introduced elevation variation coefficient to describe a terrain, which can improve the rendering efficiency of LOD algorithm when constructing a terrain mesh [20]. Relevant researchers proposed the layer-based Discrete Cosine Transform (DCT) method to simplify the terrain data in the terrain grid construction process, improving the efficiency of terrain grid construction [21]. For texture mapping, the researchers proposed a texture mapping method with multifeature control, which considered both terrain elevation value and terrain slope [22]. Researchers proposed a multiresolution texture seamless mapping method based on error control, which can realize seamless mapping by reducing projection errors based on the distance between texture block and viewpoint and the current LOD level [23]. Related scholars proved that a beach scene simulation method based on Poisson fusion, which can not only ensure realism but also effectively improve the rendering efficiency [24]. In order to build terrain mesh efficiently and improve the efficiency of terrain rendering, this paper adopts the terrain visualization method based on Levels of Detail.

3. Terrain Elevation Map Generation Based on Self-Attention Mechanism and Multifeature Sketches

To generate terrain elevation maps with rich natural features that can improve user immersion, our method first extracts terrain features based on adaptive feature enhancement to obtain terrain feature maps, then generates realistic terrain elevation maps based on self-attention mechanism, and finally realizes 3D terrain visualization based on Levels of Detail.

3.1. Terrain Feature Extraction Based on Adaptive Feature Enhancement. When generating terrain elevation maps, a large number of terrain samples and feature maps are needed as data sets. However, methods based on manual marking terrain feature maps are inefficient and cannot meet the practical application requirements. To extract terrain features efficiently and generate terrain feature maps automatically, it is necessary to study terrain feature extraction algorithms based on adaptive feature enhancement. In our method, an adaptive feature enhancement algorithm is used to preprocess the input data first. Afterwards, ridge lines, valley lines, and peak points are extracted by a profile recognition method. Finally, these terrain features can be extracted quickly and accurately as a terrain feature map.

3.1.1. Self-Adaptive Feature Enhancement. When performing adaptive feature enhancement, data need to be preprocessed first, followed by grayscale expansion. In the data preprocessing stage, when a terrain elevation map is input, the gray histogram is constructed.

As the larger the range of image grayscale, the difference between images will be more obvious, and more features can be distinguished. Hence, it is necessary to expand the

grayscale. The detailed steps of image grayscale expansion are as follows:

Step 1. Pixel number calculation: our method traverses the horizontal axis of a gray histogram in sequential order and then in reverse order to accumulate the number of pixels.

Step 2. Grayscale expansion: our method calculates the percentage of accumulated pixels to total pixels and sets a threshold K_{gray}. When this percentage reaches K_{gray}, the method records the minimum and maximum grayscale values as I_{min} and I_{max}, respectively. Pixel values that are less than I_{min} are mapped to 0 and pixel values that are greater than I_{max} are mapped to 255, as shown in (1):

$$O(i, j) = \begin{cases} 0, & I(i, j) < I_{min}, \\ 255, & I(i, j) > I_{max} \end{cases} \tag{1}$$

where $I(i, j)$ is the pixel value for row i, column j of the input image; $O(i, j)$ is the pixel value for row i, column j of the output image.

Step 3. Histogram equalization: the method performs histogram equalization for those pixels that have a grayscale value between I_{min} and I_{max} using (2), so that the pixels of each grayscale can be evenly distributed.

$$O(i, j) = 255 * \frac{I(i, j) - I_{min}}{I_{max} - I_{min}}. \tag{2}$$

The effect of grayscale expansion is shown in Figure 1, where the input image is shown in Figure 1(a), and the output image is shown in Figure 1(b). By comparison, it can be clearly seen that the terrain features of the output image are much clearer and more obvious.

3.1.2. Ridge Line Extraction. When extracting ridge lines from a terrain elevation map, we can judge whether a sampling point is a terrain feature point by comprehensively considering multiple profiles after adaptive feature enhancement. The specific process can be divided into five steps, namely, sampling map generation, feature point determination, feature-connected graph construction, ridge line prototype generation, and ridge line smoothing.

Step 1. Sampling map generation: since a terrain elevation map contains many pixels, the input data is sampled with a fixed step size to improve the efficiency of ridge line extraction. Using a method that traverses rows first and then columns, the sampling points are selected in order, and a sampling graph is generated.

Step 2. Feature point determination: in the sampling graph, the candidate points are determined by traversing the sampling points from left to right and from top to bottom. Our method traverses four profiles with the sampling point as the center and calculates the height difference between the central sampling point and other sampling points on each profile, as shown in

(a)   (b)
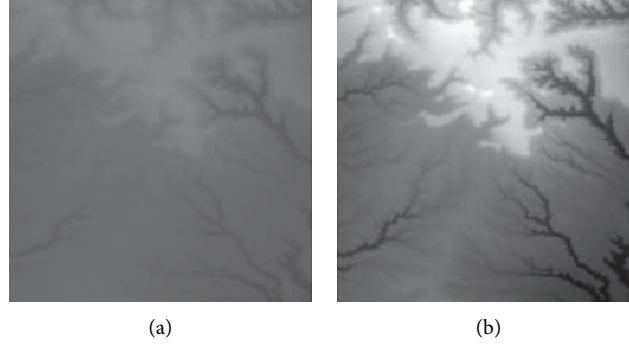
FIGURE 1: Comparisons of the effects before and after adaptive feature enhancement calculation. (a) Input image. (b) Output image.

Figure 2. For instance, in the "upper right-lower left" profile, if the elevation value of the central sampling point is higher than other sampling points on this profile, the central sampling point becomes a candidate point for the ridge line.

However, only depending on height differences to determine feature points is problematic since many ridge line feature points in valleys or with a low altitude will also be included. Hence, it is necessary to set a threshold for further screening. The threshold value is calculated based on the global maximum elevation value I_{max} and the minimum elevation value I_{min} in a terrain elevation map and uses weights to get ridge line feature points. The specific screening process is shown in (3):

$$P'_{(x,y)} \in \left\{ P_{(x,y)} | H_{(x,y)} > = w \left(I_{max} - I_{min} \right) \right\}, \qquad (3)$$

where $P_{(x,y)}$ is the previous sampling point before screening; $P_{(x,y)}$ is the sampled points after screening; $H_{(x,y)}$ is the elevation value of the sampling point; w is the weight.

Step 3. Feature-connected graph construction: based on the extracted ridge line feature points, a feature-connected graph can be constructed, which can be divided into two steps. Firstly, our method traverses the feature points from left to right and from top to bottom. Afterwards, from the current feature point, our method judges whether there are other feature points on every profile that uses the current feature point as the center. To avoid repeated calculation, only half of the profiles are traversed clockwise, and all feature points on the profile are connected, as shown in Figure 3. After traversing all the feature points, a feature-connected graph can be obtained.

Step 4. Ridge line prototype generation: based on the generated feature-connected graph, a ridge line prototype can be generated, which can be divided into three steps. Firstly, a feature edge queue is constructed. In the feature-connected graph, the edge connecting two adjacent feature points is a feature edge, and its weight is the sum of the elevation values of the adjacent feature points. Our method traverses all the feature edges of the feature-connected graph and sorts them
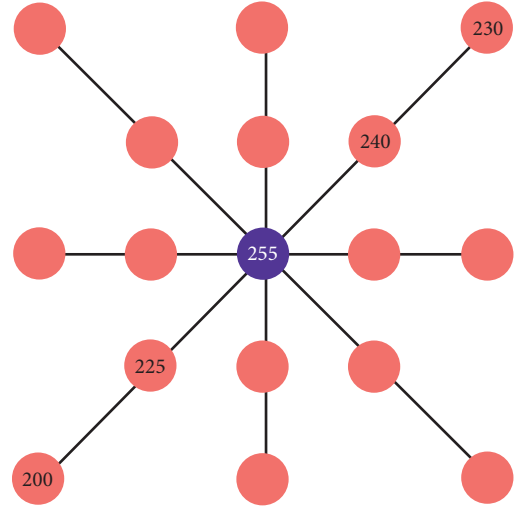
FIGURE 2: Four profiles with the current sampling point as the center.
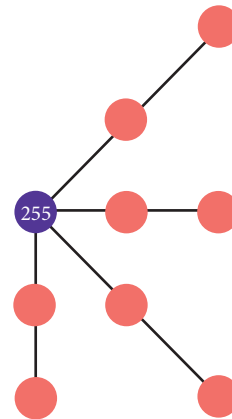
FIGURE 3: Our method traverses half of the profiles clockwise from the top right direction.

based on their weights in descending order to get a feature edge queue. Afterwards, a minimum spanning tree is generated based on the Kruskal algorithm. Finally, our method removes redundant edges by traversing the feature edges in the minimum spanning tree

and judging the number of feature edges connected by the two endpoints of the current feature edge. If the number of feature edges connected by an endpoint is less than 2, the current feature edge is redundant and can be removed. After multiple iterative removals, a ridge line prototype is generated.

Step 5. Ridge line smoothing: when smoothing ridge lines, our method takes the current feature point as the center and traverses its adjacent feature points. The method accumulates the elevation values and coordinates of all feature points and updates the position of the current feature point according to (4):

$$P'(x_0, y_0) = \frac{\sum_{i=0}^{n} H(x_i, y_i) P(x_i, y_i)}{\sum_{i=0}^{n} H(x_i, y_i)}, \tag{4}$$

where $P'(x_0, y_0)$ is the updated coordinate of a ridge feature point; $P(x_i, y_i)$ is the coordinate value of all adjacent feature points; $H(x_i, y_i)$ is the elevation value for the feature point. This operation can add some disturbance migrations to the current coordinate point, which can make the extracted ridge lines more realistic.

The process of extracting valley lines is similar to the process of extracting ridge lines. In the stage of the sampling map generation, the sampling map is inversed. The rest of the steps above can be performed to obtain the extraction results of valley lines.

3.1.3. Peak Point Extraction.

A peak point is the highest point in a local area of terrain. When dealing with peak points, our method mainly performs the extraction from two perspectives: threshold screening and profile recognition. There are two steps in extracting peak points, namely, threshold calculation and feature point determination.

Step 1. Threshold calculation: to control the number of candidate peak points, it is necessary to perform a preliminary screening of candidate points by calculating the threshold. Firstly, our method constructs a gray histogram by traversing every gray level on the horizontal axis in reverse order and accumulating the number of pixels under each gray level. When the ratio of the accumulated value to the total number of pixels is l, our method stops traversing and takes the grayscale at this moment as the threshold K_{peak}.

Step 2. Feature point determination: our method screens the pixels larger than K_{peak} in the input elevation map to get candidate points. Since the candidate points only meet the global elevation value, further screening is needed to determine the local optimal feature points, which can be divided into two steps. Firstly, our method sets a profile, which is similar to the process in ridge line extraction. Afterwards, the elevation difference on each profile is calculated clockwise.

After traversing all candidate feature points, the final extraction results of peak feature points can be obtained.

Thus, terrain feature extraction based on adaptive feature enhancement is completed. Based on the steps above, extracted ridge lines, valley lines, and peak points features can be distinguished by different colors and are drawn into a terrain feature map to obtain the final result of terrain feature extraction.

3.2. Terrain Elevation Map Generation Based on Self-Attention Mechanism.

Sketch-based terrain editing is to generate a terrain elevation map with realistic natural features based on user-edited terrain feature sketches. To generate realistic terrain elevation maps, our method applies a self-attention mechanism to the Conditional Generative Adversarial Network, realizing a terrain elevation map generation algorithm based on self-attention mechanism.

The input of the algorithm is the terrain feature map obtained in Section 3.1 or a terrain feature sketch drawn by a user. Firstly, our method constructs a generator network based on the UNet network, which is used for generating a terrain elevation map. Afterwards, our method constructs a discriminator network based on the PatchGAN network, which is used for distinguishing real terrain samples from generating results. Then, our method adds a self-attention mechanism to the generator and the discriminator, respectively, for capturing global spatial features, which can improve the performance of the generator and the accuracy of the discriminator. Finally, our method trains the discriminator and generator according to the adversarial rules to get a generator model.

3.2.1. Generator Network Construction.

In this section, we describe how our method constructs a generator based on a UNet network, whose input and output are both two-dimensional matrices so that our method can generate terrain elevation maps based on terrain feature map inputs from users. In particular, the UNet network we used adopts jumping connections, which can avoid the issue of losing input data feature details due to the increase of network depth.

The input of the generator is a terrain feature map, and the output is a terrain elevation map generated based on prediction. The whole network of the generator consists of 10 modules, which are numbered from 1 to 10, respectively. Specifically, the modules numbered from 1 to 5 consist of two convolutional layers and one pooling layer; the modules numbered from 6 to 9 consist of two convolutional layers and an upsampling layer; the module numbered 10 consists of a convolutional layer. The diagram of the network is shown in Figure 4.

The construction process of the generator network can be divided into three steps, namely, encoding unit construction, decoding unit construction, and jumping connection setup.

(1) Encoding Unit Construction. The encoding unit is used for encoding the input data to obtain the color, shape, and local spatial features from the input data. The encoding unit includes module 1 to module 5. Each module consists of two
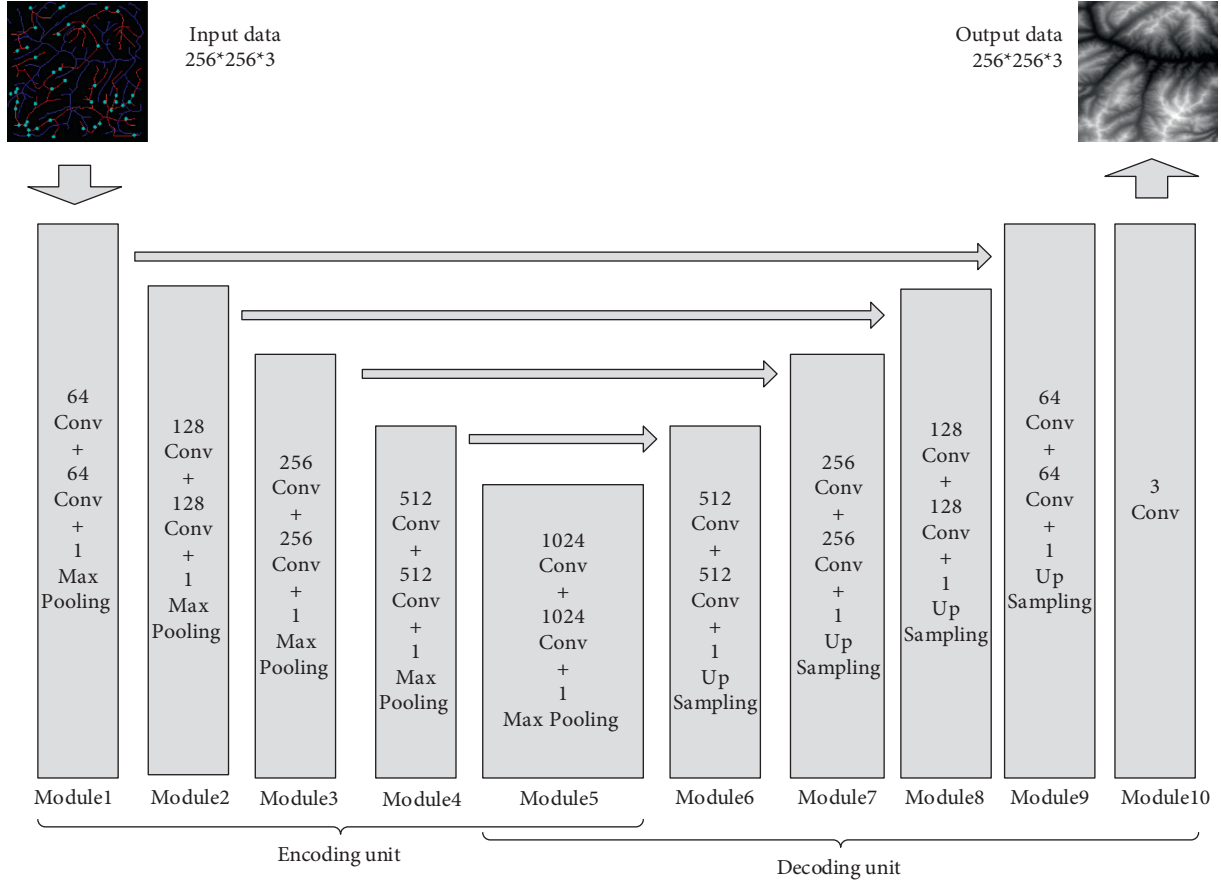
Figure 4: The structure of the generator network.

convolutional layers, one pooling layer, and one activation function. Taking module 1 as an example, the construction process includes three steps: convolutional layer construction, pooling layer construction, and activation function setup.

Step 1. Convolutional layer construction: the convolutional layer is used to encode the input and obtain the implicit features from the input terrain data. The number of convolutional layers is set to 2; the size of the convolutional kernel is set to 3×3; and the number of convolutional kernels is set to 64.

Step 2. Pooling layer construction: the pooling layer is used to compress data and reduce the number of parameters so that the overfitting issue can be mitigated. The number of pooling layers is set to 1. The pooling method adopts the maximum pooling, and the pooling layer window size is set to 2×2.

Step 3. Activation function setup: the activation function is used to activate the input and perform nonlinear output. Leaky ReLU activates the output values that are greater than 0, as shown in equation (5):

$$\text{LeakyReLU}(x) = \begin{cases} x, & x \geq 0 \\ ax, & x < 0 \end{cases} \tag{5}$$

When the input x is negative, neurons can still learn, which will not lead to deviations in the final output results.

The parameter a is generally set to 0.01. Compared with module 1, the number of convolutional kernels of each convolutional layer in modules 2–5 is twice that of the previous module, and other configurations are the same as the previous module.

(2) Decoding Unit Construction. The decoding unit is used for decoding the feature information from input and outputting the predicted result. A decoding unit includes five modules, namely, module 6 to module 10. Module 6 to module 9 adopt the combination of two convolutional layers, one upsampling layer, and one activation function. Taking module 6 as an example, the construction process includes three steps: convolutional layer construction, upsampling layer construction, and activation function setup.

Step 1. Convolutional layer construction: the number of convolutional layers is set to 2; the size of the convolutional kernel is set to 3×3; and the number of convolutional kernels is set to 512.

Step 2. Upsampling layer construction: the upsampling layer and the pooling layer are opposite in terms of functionality; that is, the values of input data are repeatedly filled in the window size area for output. The number of upsampling layers is set to 1, and the window size of the upsampling layer is set to 2×2.

Step 3. Activation function setup: similar to the encoding part, Leaky ReLU is also used for activating output, and the parameter configuration is identical.

Compared with module 6, the number of convolutional kernels of each convolutional layer in module 7 to module 9 is half of that in the previous module, and other configurations are the same as the previous module. Module 10 contains only one convolutional layer; the convolutional kernel size is set to 3×3; and the number is set to 1. The module maps the input into the output result of 3 channels and predicts the terrain elevation map.

(3) Jumping Connection Setup. Jumping connection can solve the problem that the features of shallow output are gradually lost in subsequent network with the increase of network depth. In Figure 4, the jump connection is indicated by arrows between the corresponding modules in encoding units and decoding units. Taking module 1 and module 9 as an example, the output result of module 1 is directly passed to the third dimension of module 9 input data of the decoding unit so that the shallow output features can be retained in the corresponding deep network structure. This can avoid the loss of feature information and improve the performance of the generator.

3.2.2. Discriminator Network Construction. Since Patch-GAN maps the whole input data as a probability matrix, each value in the matrix represents the discriminant result of the local input. This operation can evaluate the locally generated result and improve the performance of the discriminator. Therefore, this section describes how a discriminator based on PatchGAN is constructed, whose structure is shown in Figure 5.

There are two types of input for the discriminator: (1) real terrain features and terrain samples and (2) real terrain features and generated pseudoterrain samples. Through these two sets of data, the ability to discriminate between real results and generated results can be trained.

The construction of a discriminator network is relatively simple. Firstly, our method sets four convolutional layers, and the convolutional kernel size is set to 4×4; the number of convolutional kernels in each convolutional layer is set to 64, 128, 256, and 512, respectively. After each convolutional layer, the output is activated by Leaky ReLU. Then, our method sets a convolutional layer, where the convolutional kernel size is set to 1×1, and the number is set to 1, which is used for combining the number of input channels into 1. Afterwards, the output is activated by a Sigmoid function to obtain a binary prediction result of the local area, where the size of the output probability matrix is set to 16×16, and the number of channels is set to 1. Finally, our method accumulates all the values in the probability matrix and calculates the average value to obtain the final discrimination result.

3.2.3. Self-Attention Mechanism Construction. In the process of terrain generation, the input terrain features include not only color, shape, and local space but also global spatial features. In real terrain elevation map samples, ridge line and valley line feature areas are often continuous areas, which are all global spatial features. This section introduces a self-attention mechanism to capture the global spatial features of the input data.

The process of building a self-attention module is shown in Figure 6. The input data is the feature map output from the previous layer, and the shape of the input data is $[h, w, c]$, where the length of the input feature map is h, the width is w, and the number of channels is c. The output is the fusion result of the self-attention feature and the original feature map, whose shape is consistent with the input data. The construction process includes three steps: input data initialization, local self-attention feature map calculation, and global self-attention feature map calculation.

(1) Input Data Initialization. The input data is initialized by convolution. As shown in equation (6), our method uses convolutional kernels W_f and W_g with a number of c/k and a size of 1×1 to perform convolution operations on the input data to obtain feature maps $f(x)$ and $g(x)$. In addition, our method uses a convolutional kernel W_h with a number of c and a size of 1×1 to perform convolution operation on the input data to obtain feature maps $h(x)$.

$$f(x) = W_f x,$$
$$g(x) = W_g x, \qquad (6)$$
$$h(x) = W_h x,$$

where the shape of $f(x)$ and $g(x)$ is $[h, w, c/k]$ and the shape of $h(x)$ is $[h, w, c]$.

(2) Local Self-Attention Feature Map Calculation. Our method first readjusts the shapes of feature maps $f(x)$, $g(x)$, and $h(x)$ to $[h*w, c/k]$, $[h*w, c/k]$, and $[h*w, c]$, respectively. Afterwards, our method calculates $f(x)g(x)^T$ according to equation (7):

$$a_{j,i} = \frac{\exp(s_{i,j})}{\sum_{i=1}^{N} \exp(s_{i,j})}, s_{i,j} = f(x_i)g(x_j)^T. \qquad (7)$$

Then, the softmax function is used to activate the output of the calculation result and obtain a local self-attention feature map $a_{j,i}$ with a shape of $[h*w, h*w]$, which represents the local attention to the i-th pixel when the j-th pixel is generated.

(3) Global Self-Attention Feature Map Calculation. The global self-attention feature map is obtained by fusing the global spatial information and the local self-attention feature map, as shown in equation (8):

$$o_j = \sum_{i=1}^{h*w} a_{j,i}h(x_i), h(x_i) = W_h x_i, \qquad (8)$$

where $a_{j,i}$ represents the local self-attention feature map; $h(x_i)$ represents the global spatial information; and o_j represents the global self-attention feature map when the j-th region is generated, and the shape is $[h*w, c]$.
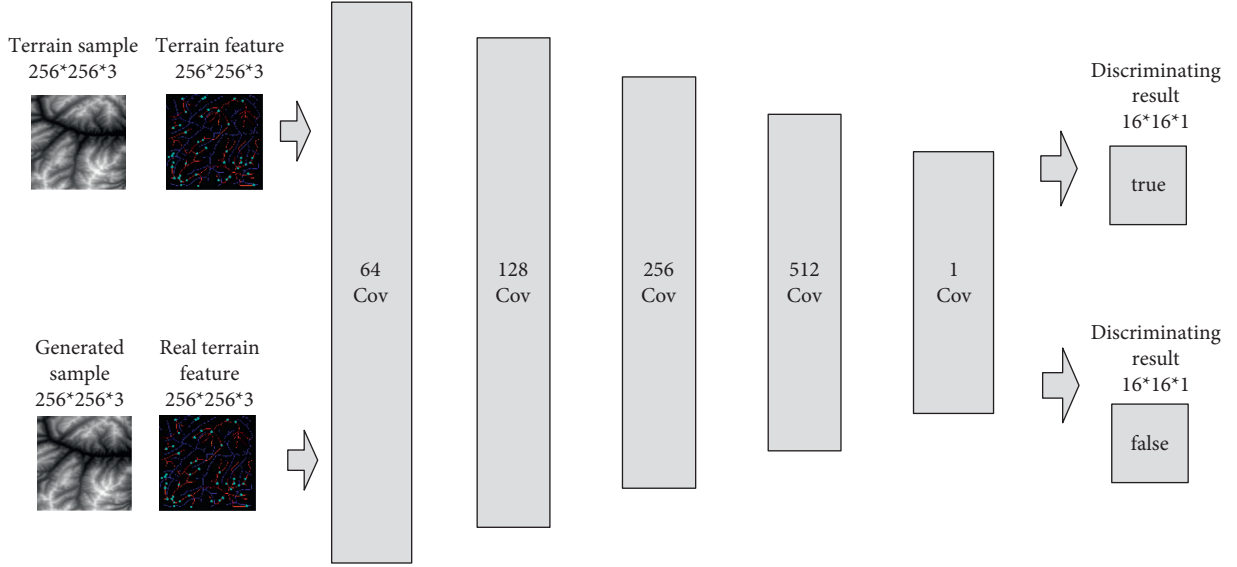
FIGURE 5: The structure of discriminator network.
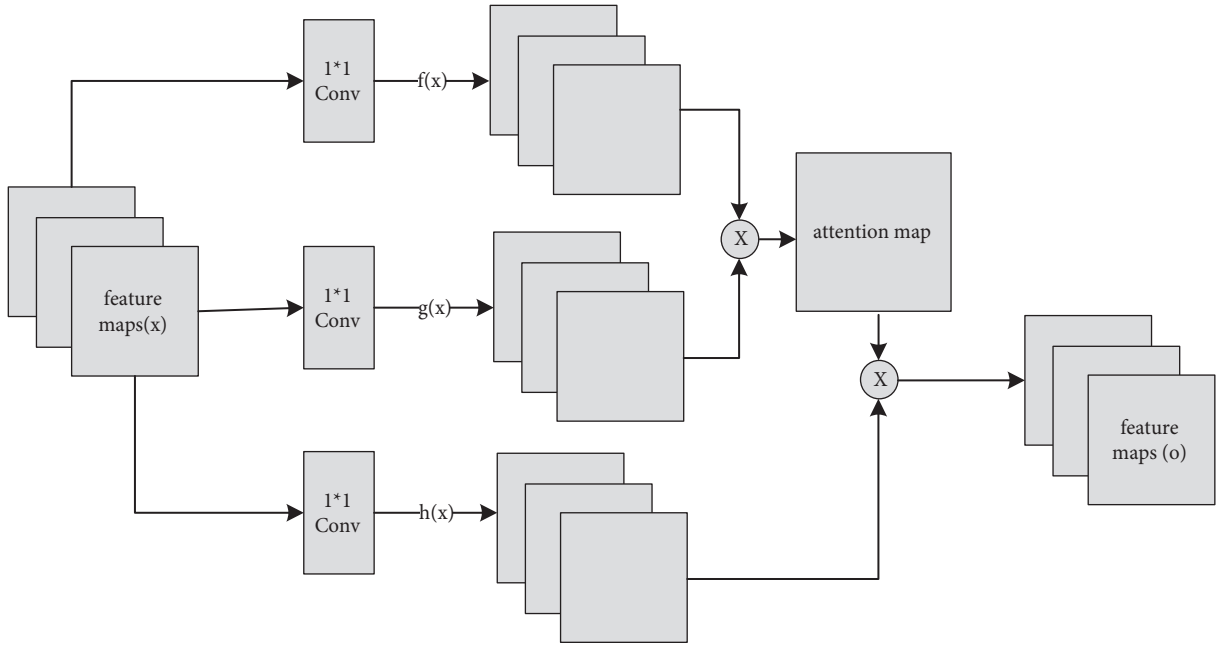
FIGURE 6: The flow of self-attention calculation.

In order to control the influence of the self-attention feature map, it is necessary to set parameters for the self-attention feature map and then fuse it with the initial input as the output. In this case, the network structure can first rely on local spatial features and then rely on remote spatial features during the training process, as shown in equation (9):

$$y_j = \gamma o_j + x_j, \tag{9}$$

where γ is the weight of self-attention with an initial value of 0. γ can be gradually assigned a larger value during the training process. The self-attention module described in this section is added after module 9 of the generator and after the

fourth convolutional layer of the discriminator, and the construction of the entire network is completed.

3.2.4. Network Training. We train the network after the network is built. The training of the network mainly includes three steps, namely, optimization target determination, training process setup, and loss function construction.

(1) Optimization Target Determination. The goal of this step is to generate a real terrain elevation map, and the evaluation standard is to judge whether it is a terrain elevation map

generated by the generator or a real terrain elevation map sample. Since this is a binary classification problem, we adopt a binary classification cross-entropy loss function to optimize the target. The calculation of the binary classification cross-entropy is shown in equation (10):

$$L_n = -\frac{1}{n} \sum_{i=1}^{n} \left[y_i \log(\widehat{y}_i) + (1 - y_i)(1 - \log(1 - \widehat{y})) \right], \quad (10)$$

where n is the number of samples; i is the current i-th sample; y_i is the label of i-th sample; $\widehat{y}_i$ is the probabilistic predicted value of i-th sample.

(2) Training Process Setup. The training process of a Generative Adversarial Network is a confrontation between generator and discriminator. The process of confrontation training is to fix the generator first, where the weights of the discriminator are trained and updated so that the ability of the discriminator to distinguish between true and false images can be maximized. Then the discriminator is fixed, where the network weights of the generator are updated so that the ability of the discriminator to distinguish between true and false images is minimized. As the generator cannot distinguish the true graph from the false graph, the performance of the generator can be improved. These processes are executed alternately until the loss function converges; the performance of the generator will be optimal at this moment, which can be used in practical applications.

(3) Loss Function Construction. According to the optimization goal and training process, the confrontation loss can be obtained, as shown in equation (11):

$$\min_G \max_D L_{G,D} = E_{(x,y)} \left[\log D(G(x \mid y)) \right] \\ + E_{(x,z)\mid y} \left[\log(1 - D(G(z \mid y))) \right], \quad (11)$$

where G is the generator; D is the discriminator; x is the real terrain data; and z is the noise data randomly generated in the generator. In addition, in order to measure the difference between the generated results and the official data at the pixel level and further improve the performance of the generator, this paper also introduces L_1 loss, as shown in equation (12):

$$L_1 = E_{(x,y)} \left[\| G(z \mid y) - G(x \mid y) \|_1 \right]. \quad (12)$$

Finally, the total loss function is obtained by synthesizing the above-mentioned confrontation loss and L_1 loss, as shown in equation (13):

$$L = L_{G,D} + L_1. \quad (13)$$

In this way, the construction of the whole network is completed. According to the training rule, we train the network weights and optimize the performance of the generator so that the trained generator can be used in the task of generating a terrain elevation map.

3.3. 3D Terrain Visualization Based on Levels of Detail. In order to build terrain mesh efficiently and ensure the smoothness of terrain visualization, this section studies the 3D terrain visualization algorithm based on Levels of Detail. Firstly, the nodes are defined by a quadtree structure, and then the terrain mesh is constructed according to the node information. Finally, multiple textures are added to the mesh.

3.3.1. Quadtree Structure Construction. Quadtree structure is a tree structure where each node includes 0 or 4 subnodes. When each node is split, the details can be improved. For a terrain mesh, when the subdivision degree of the mesh is increased, the details of the mesh will be richer. The construction of a quadtree structure can be divided into the following steps: defining nodes, calculating node coordinates, and judging whether a node needs to be divided and stored, and a complete quadtree structure can be obtained finally.

3.3.2. Terrain Mesh Construction. Based on the quadtree structure, a terrain mesh with multiple Levels of Detail can be created. Since a terrain mesh is composed of triangular faces that are surrounded by vertices, there are two steps to construct a terrain mesh, namely, creating vertices and drawing triangular faces. The drawing of triangle faces is based on the quadtree structure, where the nodes are recursively operated from the root node until all elements in the quadtree node array are traversed. Thus, a terrain mesh with Levels of Detail is constructed.

3.3.3. Multiple Texture Addition. Since 3D terrain visualization needs both terrain mesh and texture, it is necessary to add multiple textures to improve the realism of visualization results. It can be divided into three steps, namely, selecting terrain textures, creating mask textures, and fusing output. As shown in Table 1, there are four kinds of terrain textures used for visualization, namely, land, grassland 1, grassland 2, and moss, which correspond to different elevation value intervals in terrain, respectively. When performing texture mapping, the texture pixel value of a mesh vertex is calculated based on the value of mask texture. The four channel values in the mask texture and the corresponding pixel values in the four textures are multiplied and then summed, and the fused texture is the final output.

In this way, the 3D terrain visualization based on Levels of Detail is completed, which can realize the efficient construction of terrain mesh and realistic texture mapping, making interactive terrain editing more interesting.

4. Experimental Analysis

The experiment was designed to verify the feasibility and effectiveness of the multifeature sketch terrain elevation map generation method based on self-attention mechanism. The hardware environment of the verification system includes AMD Ryzen 7 4800H CPU, 16 GB RAM, and NVIDIA GeForce RTX2060 GPU. The software environment includes Windows 10 operating system, PyCharm 2019, Visual Studio

TABLE 1: Textures used in the experiment.

Land	Grassland 1	Grassland 2	Moss

2019, and Unity3D 2019. *Python* and C# are used as the main development languages.

In this study, six experiments are designed, namely, terrain feature extraction experiment, terrain generation comparative experiment, self-attention mechanism comparative experiment, terrain mesh construction comparative experiment, multiple textures addition experiment, and interactive terrain editing software experiment.

4.1. Experiment for Terrain Feature Extraction. To verify the feasibility of terrain feature extraction based on adaptive feature enhancement, a terrain feature extraction experiment was designed.

Firstly, adaptive feature enhancement was performed. Taking the terrain elevation map input in Figure 7(a) as an example, data preprocessing and grayscale expansion were performed. When expanding grayscale, K_{gray} is set to different values in Figures 7(b)–7(d). It can be seen when K_{gray} is set to 1%, the details of terrain features can be well retained while the features can be enhanced. According to the result in Figure 7(c), terrain feature extraction was performed.

The specific steps of ridge line extraction are as follows.

First, we set the sampling distance to 5 pixels to generate a sampling map. We also set the profile for the sampling points, where the length of the profile was set to 7, making preliminary screening based on the profile. Then, according to equation (3), the threshold was set for further screening. Based on our pilot study, we found that 0.03 was an ideal value for the weight w. The final result of feature point extraction is shown in Figure 8. Afterwards, the feature-connected graph was constructed by traversing all feature points, starting from the current feature point and connecting the feature points on the profile. The final construction result is shown in Figure 9. Then, we sorted the feature edges and stored them in the feature edge queue to create a minimum spanning tree according to the feature edge queue, as shown in Figure 10. The number of iterative deletions was set to 3, so the redundant edges in the minimum spanning tree can be removed to obtain a prototype of the ridge feature line, as shown in Figure 11. Then, smoothing was performed on the feature-connected graph to make the feature lines more similar to real terrain features. The smoothing result is shown in Figure 12.

The valley line extraction process is similar to the ridge line extraction process. When generating a valley line sampling map, the sampling map is inversed, and then the steps above are performed to obtain the valley line extraction result, as shown in Figure 13.

Finally, we extracted the peak points. First, we calculated the threshold value. Our pilot study showed that when the accumulated number of pixels accounted for 5% of the total number of pixels, the number of selected peak points was more suitable. Then, we determined the feature points and the peak points through the profile recognition and elevation difference. The extraction result of the peak point is shown in Figure 14.

4.2. Comparative Experiment for Terrain Generation. To verify the feasibility of terrain elevation map generation based on a self-attention mechanism, this experiment was performed from two aspects: (1) using real terrain features to demonstrate generated terrain elevation maps and compare the differences between the generated results and real terrain samples; (2) using hand-painted terrain features to demonstrate generated terrain elevation maps and 3D visualization effects.

(1) Using Real Terrain Features. We input real terrain features in the trained generator model for prediction and compared the differences between the generated terrain elevation maps and real terrain samples.

The experimental results are shown in Table 2, where the features of ridge lines, valley lines, and peak points in the generated terrain elevation maps were consistent with the features in real terrain feature maps. To evaluate the similarity between the generated results and the real terrain results, the SSIM image similarity evaluation algorithm was used. The calculation result of SSIM is between 0 and 1; the closer the result is towards 1, the more similar the two images will be. It can be seen that the SSIM results obtained by our method were all greater than 0.5, which suggested that the similarity with the original image was high. This demonstrated the feasibility of our proposed method, which can meet the needs of interactive terrain editing.

(2) Using Hand-Painted Terrain Features. To further verify the feasibility of our proposed method, this experiment input terrain feature sketches drawn by users into the trained

(a)   (b)   (c)   (d)
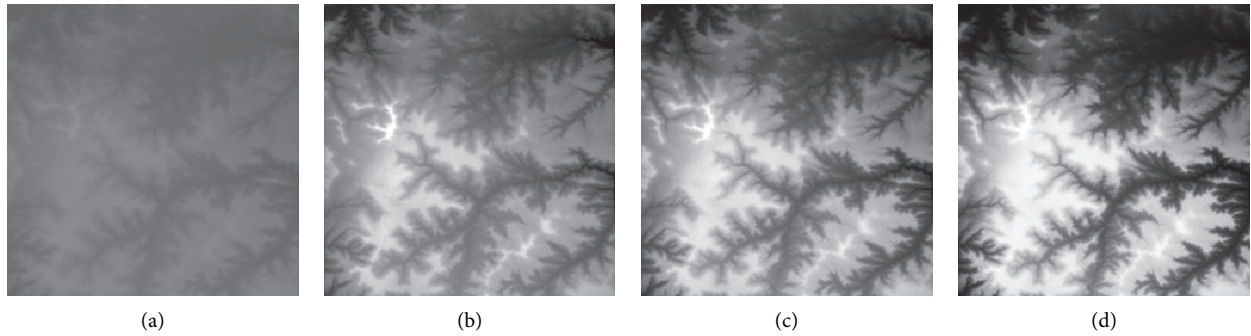
FIGURE 7: Comparison of original input data and enhanced data. (a) Input of a terrain elevation map. (b) Result when K_{gray} is set to 0.05%. (c) Result when K_{gray} is set to 1%. (d) Result when K_{gray} is set to 1.5%.
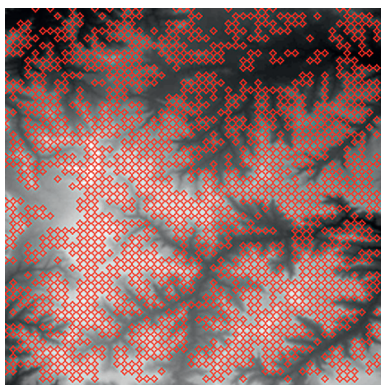
FIGURE 8: The candidate feature points for ridge lines obtained by profile recognition and screening.
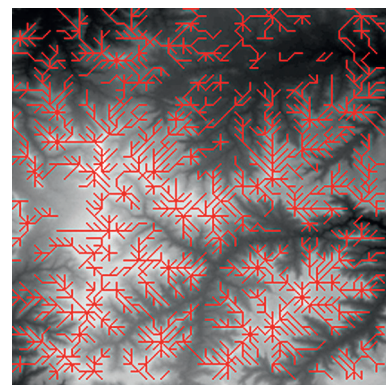
FIGURE 10: The minimum spanning tree extracted from feature-connected graph.
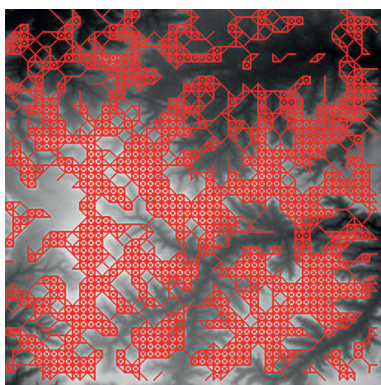
FIGURE 9: Feature-connected graph build based on feature points.
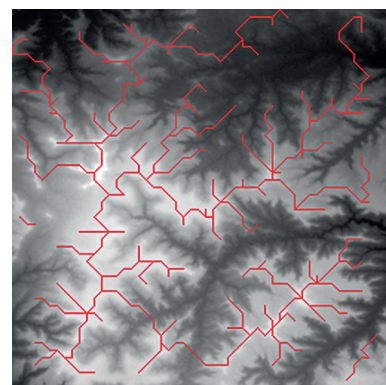
FIGURE 11: Feature line prototype after redundant edge removal.

generator model for prediction and demonstrated the generated terrain elevation maps and 3D visualization results.

The experimental results are shown in Table 3. It can be seen that terrain feature sketches can be very simple, which only have a few simple strokes. We input these sketches into the trained generator model, which can generate highly realistic terrain elevation maps based on these sketches.

We also used World Machine software to render the terrain elevation maps based on user sketches to further verify the realism of editing results. It can be seen that the rendered terrain results have realistic ridge lines, valley lines, and peak points, which can meet the demands of interactive editing.

4.3. Comparative Experiment for Self-Attention Mechanism. To verify the effectiveness of the self-attention mechanism in terrain elevation map generation, this experiment compared the results with and without the self-attention mechanism.
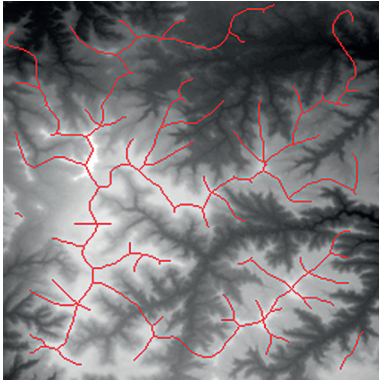
FIGURE 12: The final result of ridge line feature extraction.
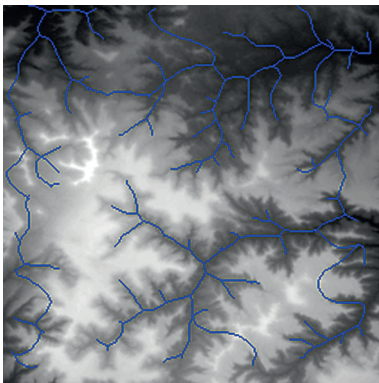
FIGURE 13: The result of valley line feature extraction.
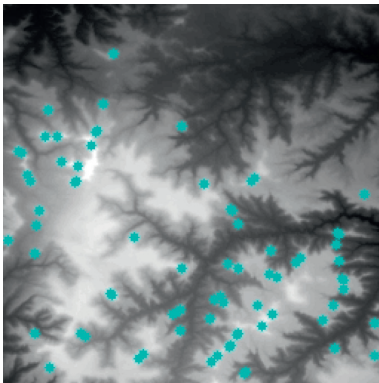
FIGURE 14: The result of peak point feature extraction.

The input was the terrain features of real samples, as shown in Table 4. It can be seen that some ridge line and valley line structures in the generated results are incomplete without a self-attention mechanism, as shown in the places marked by red circles. In contrast, the results with the self-attention mechanism are much better in terms of completeness and detail, and the generator model also has a higher performance.

4.4. Comparative Experiment for Terrain Mesh Construction. To verify the feasibility and effectiveness of the terrain mesh construction in our proposed method, a terrain mesh

construction experiment based on Levels of Detail was designed.

First, we established a quadtree structure. Afterwards, we constructed a terrain mesh and a vertex array. According to whether each node can be divided, a terrain mesh was constructed by drawing triangle faces. In the scene view in Unity3D editor, WireFrame mode was used to display the terrain mesh. The result is shown in Figure 15.

It can be seen that the terrain mesh constructed based on Levels of Detail has lower mesh complexity. While the areas that are far from the viewpoint have lower details, the areas that are close to the viewpoint have greater details. By moving the viewpoint, the changes of details can be clearly seen in local terrain areas. As shown in Table 5, when the camera moved towards the terrain mesh, the terrain details became gradually enriched. Thus, our method realized dynamic control of mesh details, and rendering efficiency was high.

4.5. Experiment for Multiple Texture Mapping. To verify the effectiveness of using multiple textures to enhance the realism of 3D terrain, an experiment of multiple texture mapping was designed.
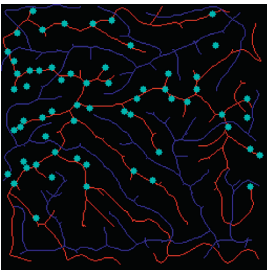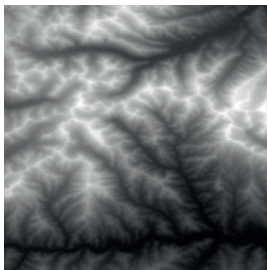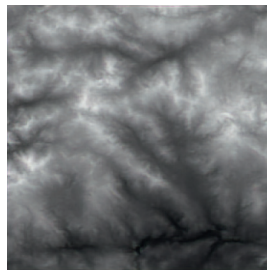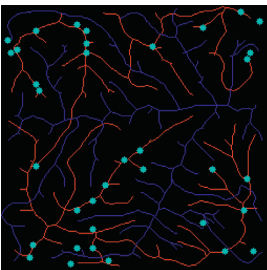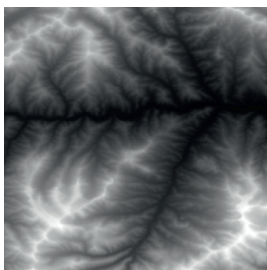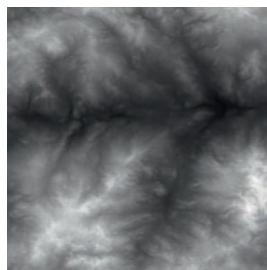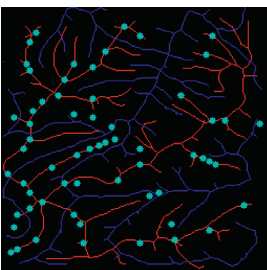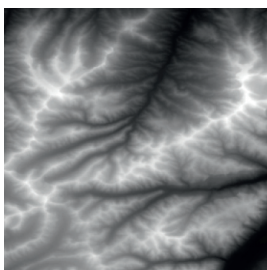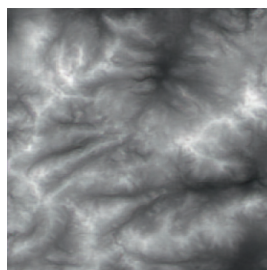
First, a mask texture was created based on a terrain elevation map drawn by a user. Then, the sample texture and mask texture were fused for outputting multiple textures, which adopted different elevation value intervals for different texture mapping. Finally, random translations and mirror operations were performed on the sample texture to reduce texture repetitions during the tiling process.

The experimental results are shown in Figure 16. It can be seen that multiple texture mapping can assign different textures based on the elevation value intervals of the terrain elevation map. From 3D visualizations of terrains, users can have an intuitive and fun experience in interactive terrain editing and terrain elevation map creation.

4.6. Experiment for Interactive Terrain Editing. To provide a good user experience for interactive terrain editing, we designed an interactive terrain editing tool based on the methods for terrain feature extraction, terrain elevation map generation, and 3D terrain visualization that are described earlier in this paper. With the help of this tool, users can freely draw ridge line, valley line, and peak point features on a sketch. The tool can generate terrain elevation maps with realistic features in real time based on the sketch and provide 3D visualization for the terrain that users can interact with.

The interactive terrain editing tool based on a multi-feature sketch can be divided into four modules: terrain feature extraction module, terrain elevation map generation module, terrain visualization module, and human-computer interaction module. Specifically, terrain feature extraction is used for extracting terrain features, creating data sets, and training networks to get the generator model with the best performance. With this generator model, the terrain elevation map generation module can generate a terrain elevation map based on a terrain feature sketch drawn by a user. The terrain visualization module adopts LOD mode to build

TABLE 2: Generated terrain results based on real terrain features.

Terrain features	Real terrain samples	Generated results	SSIM
			0.60
			0.65
			0.63

a terrain mesh and adopts multiple texture mapping to improve the realism of 3D terrain. Finally, the human-computer interaction module provides a user interface for the functionalities above, creating a good user experience for the interactive terrain editing tool.

(1) Terrain Feature Extraction Module. Terrain feature extraction module is the core module for interactive terrain editing. In the terrain editing tool, terrain features are distinguished by using different colors and drawing methods. Specifically, red lines are used for ridge lines; blue lines are used for valley lines; cyan dots are used for peak points.

(2) Terrain Elevation Map Generation Module. Terrain elevation map generation module is the basis for real-time interactive terrain editing. The trained generator model can be loaded by the LoadModel() method in the Keras library. A series of format conversions are performed on the user-created feature sketch to obtain a suitable data format for the model input, and the Predict() method is called to predict and generate a terrain elevation map. Finally, the generated

result is converted to PNG format, where the user can choose different resolutions to save the file.

(3) Terrain Visualization Module. Terrain visualization module allows users to observe the generated results intuitively. The module includes two submodules: mesh construction and texture mapping. Due to the LOD-based mesh construction, it is possible to ensure the user experience on low-performance devices.

(4) Human-Computer Interaction Module. Human-computer interaction module can have a great impact on user experience for the terrain editing tool. The module can be divided into two submodules: interactive user interface module and interactive function module. Specifically, the interactive user interface module includes a main user interface, a feature extraction user interface, and a roaming user interface; interactive function module includes terrain feature extraction function, editing and saving function, and roaming display function.

The initial user interface of our interactive terrain editing tool is shown in Figure 17. The black area on the

TABLE 3: Generated terrain results based on sketch input.
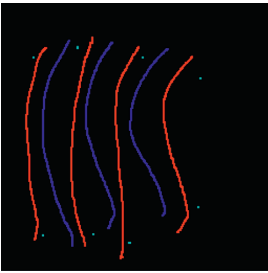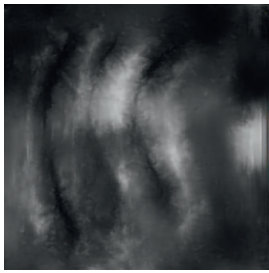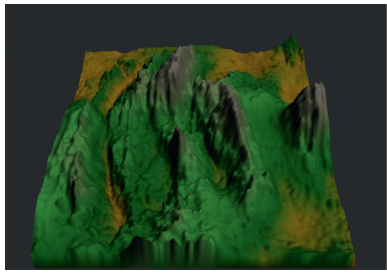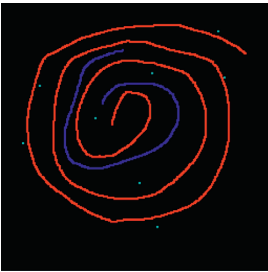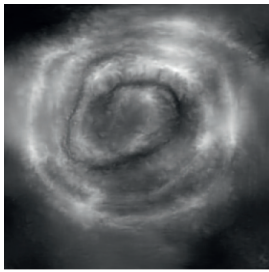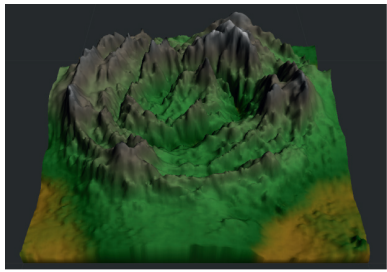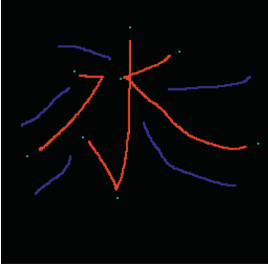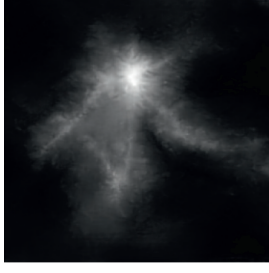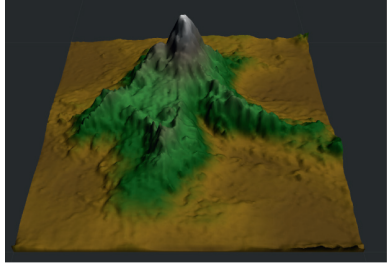
Terrain feature sketches	Generated results	Rendered results

TABLE 4: Comparison of generated terrain elevation maps with and without self-attention mechanism.
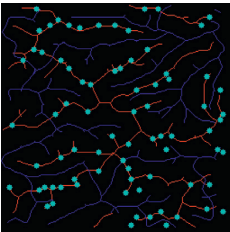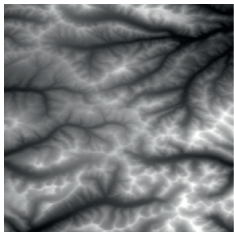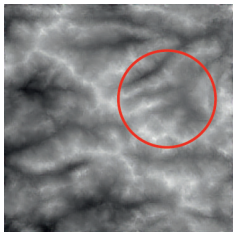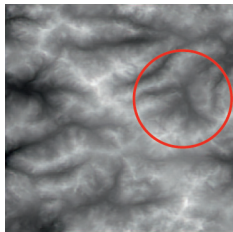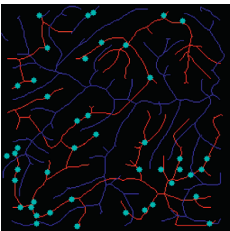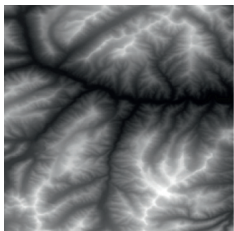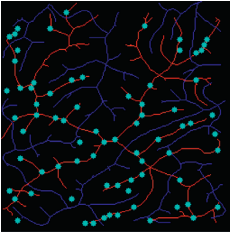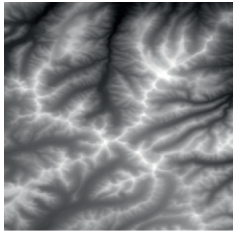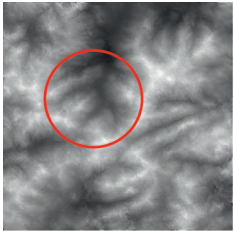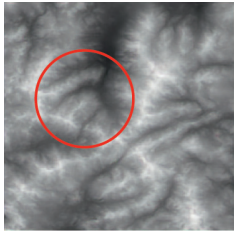
Terrain features	Real terrains	Without self-attention mechanism	With self-attention mechanism

TABLE 4: Continued.

Terrain features	Real terrains	Without self-attention mechanism	With self-attention mechanism

FIGURE 15: The result of terrain mesh constructed based on Levels of Detail.

TABLE 5: Generated terrain meshes based on different distances.

Distance = 9000	Distance = 6000
Distance = 3000	Distance = 1000
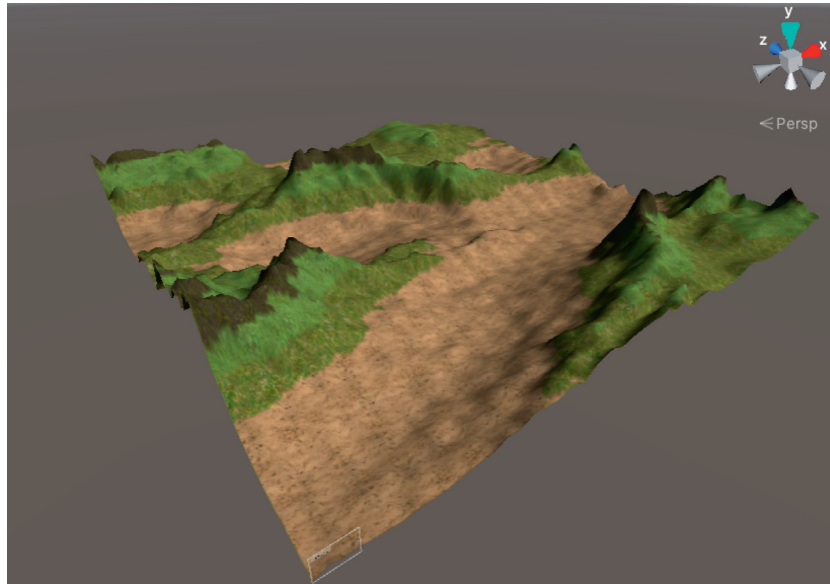
Figure 16: The result of adaptive texture mapping.
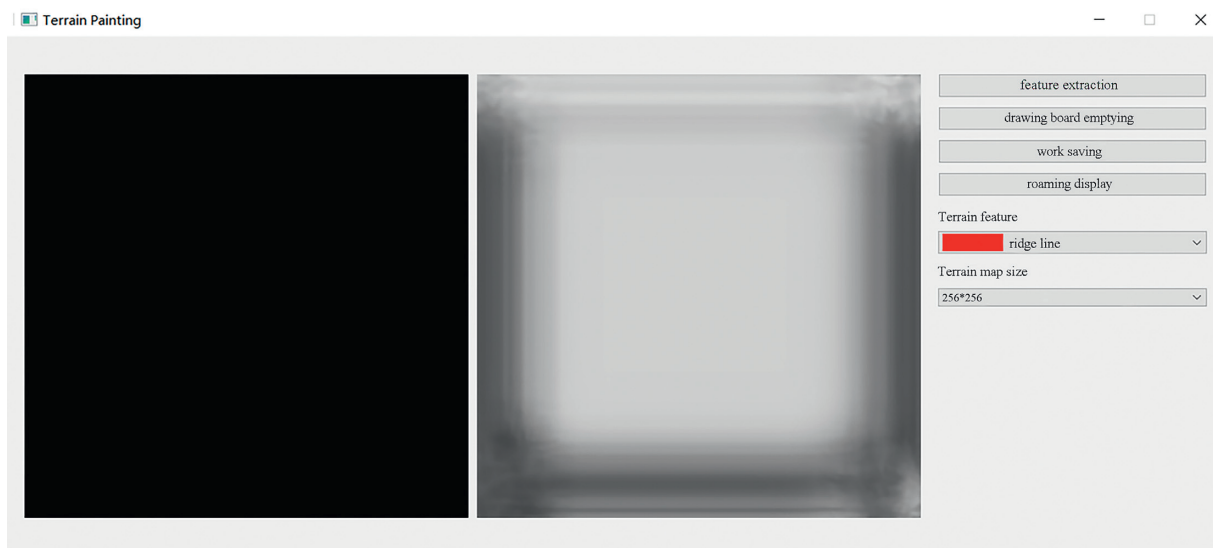
Figure 17: The initial user interface of our interactive terrain editing tool.
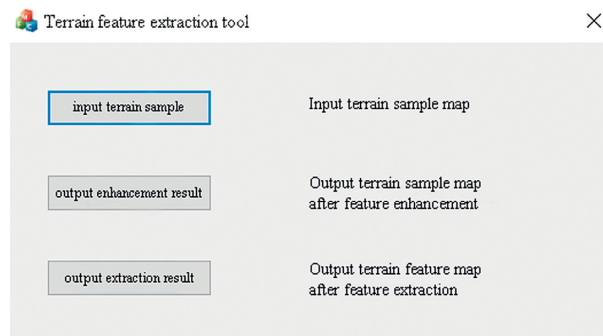
Figure 18: The interface of terrain feature extraction.
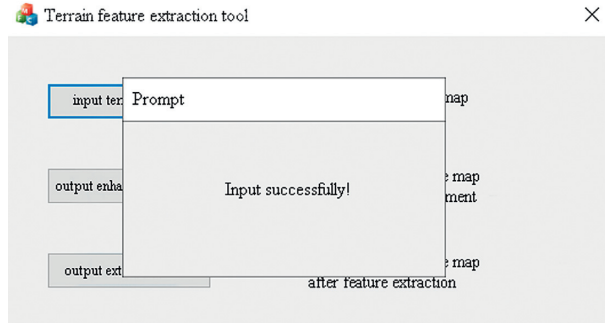
FIGURE 19: The interface for successful input.
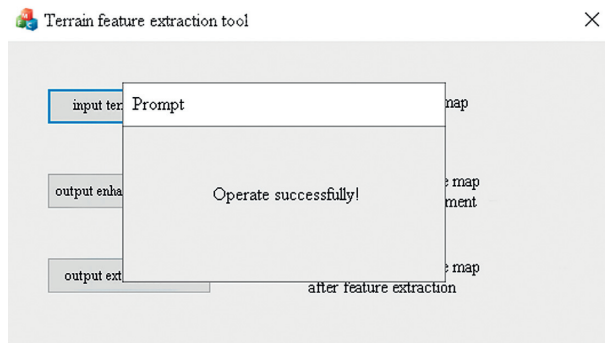
FIGURE 20: The interface for successful feature enhancement and feature extraction.
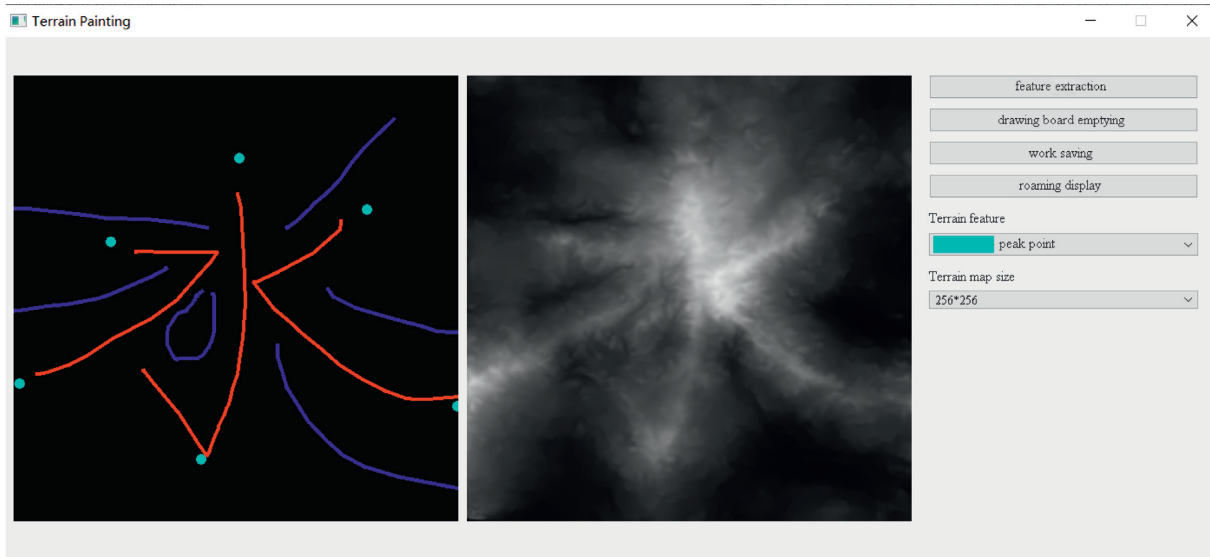
FIGURE 21: The generated terrain elevation map based on a user sketch.

left side is the user sketch drawing area, and the white area on the right side is the terrain generation area. The rightmost sidebar has six functions, namely, feature extraction, drawing board emptying, work saving, roaming display, terrain feature selection, and output resolution selection.

Users can click the feature extraction button to start terrain feature extraction, which can perform feature enhancement and feature extraction operations on the local terrain elevation map data sets. The user interface for terrain feature extraction is shown in Figure 18. After successful input, feature enhancement, and extraction, a popup will be

FIGURE 22: The result of snow-style terrain visualization.

displayed to indicate the operation is successful, as shown in Figure 19 and Figure 20.

Users can select different brushes to draw terrain features on a sketch, and the generated results will be displayed in the terrain generation display area. As shown in Figure 21, three types of features, including ridge lines, valley lines, and peak points, are drawn on the sketch. Since the low-level GPU-driven generator is used, the calculation can be done in real time. When a user is drawing terrain features on the sketch, the prediction of terrain elevation map can be completed in real time.

After drawing, the user can click the roaming display button to start interactive roaming to observe the visualization effect of 3D terrain. To increase the fun for user interaction, users can click the number key 1 to display a snow scenery, which can display a terrain under different weather conditions, as shown in Figure 22.

5. Conclusions and Future Work

To address the issues of low efficiency in obtaining terrain feature map by manual annotation and poor realism in generated terrain elevation map, this paper proposes a method of generating terrain elevation map based on self-attention mechanism and multifeature sketch. First, a terrain feature extraction method based on adaptive feature enhancement and profile recognition is proposed to realize the rapid and automatic terrain feature extraction of ridge lines, valley lines, and peak points, which can perform well in large-scale terrain feature map generations. Afterwards, our method adopts terrain elevation map generation based on the self-attention mechanism, where a generator network based on UNet network and a discriminator network based on PatchGAN network are built, and then self-attention mechanism are added to the generator and the discriminator, respectively, to capture the global spatial features for generating a realistic terrain elevation map. Finally, by constructing a terrain mesh and adding multiple textures, a 3D terrain visualization method based on Levels of Detail is proposed, which can provide high rendering efficiency and realistic visualization results. In this paper, an interactive terrain editing tool based on a multifeature sketch is implemented, which allows users to interact with the generated terrain. Experimental data show that the proposed method is effective in terrain feature extraction, terrain elevation map generation, and 3D visualization. In addition, the editing tool can run smoothly, providing intuitive user interactions and a good user experience.

Since the research on interactive terrain editing techniques is conducted under specific conditions, there are certain limitations and spaces for optimization. Here are a few points that can be further studied in theory and practice:

(1) *Network Structure Optimization.* The network adopted in this paper is based on the Conditional Generative Adversarial Network structure. In the future, we can explore the influence of different network structures on the generation results and the efficiency of different training models and further optimize the network structure.

(2) *Texture Mapping Optimization.* Despite the fact that multiple texture mapping is used in this paper, there is a limit on the number of textures. In the future, we can explore new ways to create mask textures, eliminating the limitation on the number of sample textures to achieve more realistic texture mapping.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] H. C. Song, Y. Gao, and Y. M. Wei, "Interactive generation and representation of three-dimensional terrain," *Computer Engineering and Applications*, vol. 31, pp. 80–82, 2004.

[2] H. F. Yin, C. W. Zheng, and X. H. Hu, "Interactive digital terrain synthesis algorithm," *Journal of Computer-Aided Design and Graphics*, vol. 24, no. 7, pp. 909–917, 2012.

[3] Y. Y. Zeng, F. J. Kang, and H. Z. Yang, "Adaptive multi-feature fusion for fast realistic terrain mapping," " *Journal of Image and Graphics*, vol. 18, no. 6, pp. 724–729, 2013.

[4] Y. Liu, X. G. Tou, and H. L. Li, "An efficient parallel method for topography generation by Perlin noise," *Bulletin of Science and Technology*, vol. 32, no. 3, pp. 200–204, 2016.

[5] É. Guérin, J. Digne, E. Galin et al., "Interactive example-based terrain authoring with conditional generative adversarial networks," *ACM Transactions on Graphics*, vol. 36, no. 6, pp. 1–13, 2017.

[6] G. Cordonnier, M.-P. Cani, B. Benes, J. Braun, and E. Galin, "Sculpting mountains: interactive terrain modeling based on subsurface geology," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 5, pp. 1756–1769, 2018.

[7] A. W. Jensen, N. N. Rant, T. N. Møller, and J. A. Billeskov, "Deep convolutional generative adversarial network for procedural 3D landscape generation based on DEM," in *Proceedings of the 6th EAI International Conference on Interactivity and Game Creation*, pp. 85–94, Heraklion, Greece, October 2017.

[8] A. Paris, E. Galin, A. Peytavie, E. Guérin, and J. Gain, "Terrain amplification with implicit 3D features," *ACM Transactions on Graphics*, vol. 38, no. 5, pp. 1–15, 2019.

[9] R. Fischer, P. Dittmann, R. Weller, and G. Zachmann, "AutoBiomes: procedural generation of multi-biome land-scapes," *The Visual Computer*, vol. 36, no. 10-12, pp. 2263–2272, 2020.

[10] Y.-C. Chang, G.-S. Song, and S.-K. Hsu, "Automatic extraction of ridge and valley axes using the profile recognition and polygon-breaking algorithm," *Computers & Geosciences*, vol. 24, no. 1, pp. 83–93, 1998.

[11] P. Huang and Z. Liu, "Extraction of ridge and valley lines based on terrain gradient direction," *Journal of Wuhan University (Natural Science Edition)*, vol. 5, pp. 396–399, 2005.

[12] H. Zhang, Y. Liu, and Z. Ma, "A terrain skeleton feature extraction method based on morphological coding," *Computer Research and Development*, vol. 52, no. 6, pp. 1409–1423, 2015.

[13] K. Zou, H. Wong, and W. Li, "DEM terrain feature line extraction with controllable saliency," *Journal of Image and Graphics*, vol. 22, no. 11, pp. 1611–1622, 2017.

[14] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial networks," in *Proceedings of the Neural Information Processing Systems Conference (NIPS)*, vol. 3, pp. 2672–2680, Montreal, Quebec, Canada, December 2014.

[15] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *Computer Science*, vol. 22, pp. 2672–2680, 2014.

[16] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proceedings of the International Conference on Learning Representations (ICLR)*, San Juan, Puerto Rico, May 2016.

[17] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition (CVPR)*, no. 5967-5976, Las Vegas, NV, USA, June 2016.

[18] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 12744–12753, Stockholm, Sweden, July 2018.

[19] J. H. Clark and H. James, "Hierarchical geometric models for visible surface algorithms," *Communications of the ACM*, vol. 19, no. 10, pp. 547–554, 1976.

[20] Z. Wang and X. Lü, "Terrain rendering LOD algorithm based on improved restrictive quadtree segmentation and variation coefficient of elevation," *Journal of Beijing Institute of Technology (Social Sciences Edition)*, vol. 27, no. 4, pp. 145–150, 2018.

[21] D. Du, B. L. Gao, and L. Tian, "Combining fast layer-based DCT and dynamic LOD for terrain compression mapping techniques," *Computer Engineering and Applications*, vol. 56, no. 13, pp. 223–229, 2020.

[22] R. Zhao, Y. Zhang, and J. Wang, "Multi-map terrain texture synthesis algorithm under feature control," *Computer Engineering*, vol. 38, no. 10, pp. 197–199, 2012.

[23] B. Gao, B. Zhang, and M. Dou, "Multi-resolution texture seamless mapping based on error control," *Computer Engineering and Design*, vol. 39, no. 12, pp. 3774–3778, 2018.

[24] Y. Zhou, J. Li, and W. Xu, "A real-time beach scene simulation based on Poisson fusion," *Journal of Wuhan University (Natural Science Edition)*, vol. 51, no. 4, pp. 363–370, 2018.

Hindawi

Research Article

Research on the Prediction of Popularity of News Dissemination Public Opinion Based on Data Mining

Jing Tian [iD], Huayin Fan, and Zengwen Hou

Journalism & Communications College, Jilin Normal University, Changchun 130012, Jilin, China

Correspondence should be addressed to Jing Tian; tj80@jlnu.edu.cn

The development of news public opinion presents the characteristics of dynamic changes, and its life cycle is generally relatively short. For news public opinion to be welcomed by everyone, that is, to become hot news, it must be able to spread to a large number of readers in a short time. And, some of its characteristic attributes must satisfy the interests of most users and arouse users' desire to read. Therefore, it is particularly important to extract and study these characteristic attributes that determine the popularity of news public opinion and finally establish a model to describe the relationship between the popularity of news public opinion and these characteristic attributes. Based on data mining, this article mainly studies the popularity of news public opinion from two aspects. First is the sampling and attribute feature extraction of news. Then, considering the nonlinear relationship between the features, an improved principal component analysis method is proposed to analyze the correlation of the features. This can select important features from many irrelevant features and effectively reduce the original high-dimensional features. Second, the application of neural network in the prediction of news public opinion is studied. BP is an efficient data mining method. Considering BP network has some shortcomings. This work uses an improved particle swarm optimization to optimize the initial parameters for BP network, which can compensate for the defects of BP network. After that, the BP network with optimized parameters is used to establish a prediction model for the popularity of news public opinion. The experimental results prove that the neural network model proposed can accurately predict the popularity of news public opinion.

1. Introduction

The term news was first born in the Tang Dynasty and refers to the time that was recently heard or what happened recently in society. The meaning of the word news has not changed much after more than a thousand years of development. However, the dissemination of news has undergone earth-shaking changes after more than a thousand years of historical changes. Traditional news dissemination, from small-scale oral transmission to newspapers, to wireless broadcasting, and to TV broadcasting, blooms in many ways and has various channels. Nowadays, the news is presented on everyone's mobile terminal, and traditional news dissemination methods have gradually withdrawn from the market [1–5].

With the rapid development of the Internet, the Internet has become a part of people's lives. People use mobile phones to receive information and view news anytime and anywhere on the Internet. Compared with traditional news dissemination, the Internet not only brings more traffic but also shortens the generation, creation, and dissemination time of news exponentially. The data shows that more than 2,600,000 news are shared every minute on Facebook, the average number of online videos shared every minute on LINE is 8,333, and the number of news links shared every minute on Twitter is close to 300,000. Interconnection has detonated the information age, and news hotspots have been generated all the time. Traditional news dissemination is restricted by channels, and the audience is limited. Even if it is a hot topic, its influence is also limited. But nowadays, the coverage of breaking news is so wide that it is unimaginable [6–10].

In the Internet age, news and public opinions cover all aspects of our lives. The real-time, uncertain, and rapid

spread of the Internet make the characteristics of online news public opinion show the characteristics of the new era such as randomness, suddenness, and concealment. In such an era, everyone on the Internet can become a publisher and disseminator of information, as well as a creator of news and public opinion. Of course, news and public opinion are not all good, and the domestic Internet environment is not a pure land, which is full of negative, distorted, extreme, inciting violence and division factors. Popular news, especially breaking news, as a powerful means of communication, is not only conducive to the rapid dissemination of information but also conducive to the transfer of value. Under such a background, predictive analysis of the popularity of news and public opinion is beneficial to the healthy development of the Internet and curbs negative information such as rumors and violence, at the same time, producing news or advertisements that can arouse people's general goodwill. It has great value for many Internet publishers, portal websites, and Internet advertising companies. The development of news public opinion presents the characteristics of dynamic changes, and its life cycle is generally relatively short. If a piece of news is to be welcomed by everyone, it becomes hot news. It must be able to spread to a large number of readers in a relatively short period of time, and some of its characteristic attributes must satisfy the interests of most users and arouse users' desire to read. Therefore, it is particularly important to extract and study these characteristic attributes that determine the popularity of online news, analyze their interrelationships, and finally, establish a model to describe the relationship between the popularity of news public opinion and these characteristic attributes [11–15].

The contributions are as follows: (1) This work proposes an improved principal component analysis method to analyze the correlation of features. In this way, important features can be selected from many irrelevant features, effectively reducing the dimensionality of the original high-dimensional features. (2) This work uses an improved particle swarm optimization to optimize the initial parameters for the BP network. It can make up for the shortcomings of BP and use the BP network with optimized parameters to establish a prediction model for the popularity of news and public opinion.

2. Related Work

Compared with traditional news and advertising industries, Internet news, with its low-cost, fast-spreading, and wide-ranging features, had a huge impact on the traditional news industry and had shown initial advantages. Due to the real-time characteristics of news and the guiding characteristics of public opinion, accurately predicting the popularity of a news manuscript for the roles of journalists, content providers, advertisers, etc., had become a new research hotspot. However, predicting the popularity of news articles on the Internet was a challenging task. First of all, the content of these articles was only available online. The structure of the social network that spreads news was different, which would affect the probability that users can see this article. In addition, the coverage of the content of the article was also a

factor that affects the prediction of the results. For example, articles that were closely related to the general public on a global scale are more popular than articles that only targeted a small number of social groups.

For this challenging work, a lot of research had been devoted to it, and a lot of questions had been raised. Literature [16] put forward the problem of tracking the spread of online topics. Literature [17] proposed detecting the public opinion orientation of online communities. The literature [18] put forward the field of prediction of the popularity of social media. There were many different ways of expressing the concept of popularity. For example, our most common classic method was the number of clicks on a piece of news, but this information was of little significance to researchers. Because in fact, due to the existence of web crawlers and search engines, it was difficult for us to estimate the number of times, an article was actually requested by users, which would directly lead to the inaccuracy of the original data. In addition, there were some other methods based on the user's own participation, such as comments, voting, and sharing via email or WeChat. Predicting the popularity of news articles was a difficult task, and a lot of research results had been accumulated over the years. Literature [19] proposed a model to predict the popularity of news published by science and technology information websites. This method believed that the popularity of articles will gradually increase over time. Literature [20] proposed two algorithms for predicting video popularity and news popularity and achieved good prediction results. Literature [21] proposed a linear logarithm method, which can meet the forecasting demand in a specific range. Literature [22] proposed a different way of thinking. He established a new prediction model based on the social influence of news and the characteristics of the publishing platform. The disadvantage of this type of algorithm was that it could only be applied to predictive analysis in a specific field but failed to analyze articles on different topics. Literature [23] proposed a new forecasting idea, which focused on the prediction of the popularity of certain news that would continue to be paid attention to after a certain period of time. Literature [24] proposed a method. Using the number of news sharing as a popular indicator, a clustering algorithm was used to divide different articles into 4 categories according to their popularity. Literature [25] proposed a new idea to predict the popularity of online news and ranked it according to the popularity. The disadvantage of this type of method was that it could not give an accurate prediction value.

Driven by the Internet industry chain, Internet + had become an industrialized concept, and network news was quickly spread through the Internet as a carrier. Data showed that more than 26,000 news were shared per minute on Facebook, the average number of online videos shared per minute on LINE was 8333, and the number of news links shared per minute on Twitter was close to 3,000. Clarifying how popular the public was for a piece of news has become a new branch of research.

At present, many models for tracking and forecasting online news based on time series had been proposed. These models assumed that the evaluation of popularity was time-

sensitive. The number of users of YouTube videos fluctuated continuously within 24 hours, and the number of their shares was also constantly changing. The literature [26] aimed at analyzing the life cycle of online news; this research used a threshold value to compare and calculate the number of news shares in 7 days. At the same time, the feature extraction method was another subtle branch in the field of social news research. It mainly focused on the news itself and removed some possible influence factors, such as time, comment, and emotional color. The reason why time was eliminated was that the number of news sharing gradually increases with the passage of time. But its impact on users would gradually become zero after a certain point in time. In addition, the results of this research field were diversified because of the existence of various media forms, including videos, music records, and plain text information [27].

3. Method

3.1. Improved Principal Component Analysis

3.1.1. Algorithm Background. When predicting the popularity of news public opinion, it is necessary to extract the characteristics of the news. Under normal circumstances, these features are high-dimensional, and processing these features is more complicated. Therefore, it is necessary to reduce the dimensionality of the features before making subsequent predictions. Principal component analysis has become a classic and effective statistical dimensionality reduction method. It has been successfully applied to linearly distributed data for a variety of practical problems and has also achieved remarkable results. At the same time, the algorithm for further expansion of principal component analysis has also been effectively developed.

First of all, through research, we can know that sparse principal component analysis is based on principal component analysis, which mainly transforms the problem of solving the principal component load into a ridge regression problem. A regularization penalty term is added to the essence of such an optimization regression problem to compare the load of each principal component. This promotes a load of some relatively weak variables to approach zero, achieving a sparse effect. At present, there are generally two types of penalty items in sparse principal component analysis: L_1 penalty items and TL_1 penalty items. The L_1 penalty term was originally proposed by the literature [28] when the L_1 penalty term was combined with the ridge regression term in the optimization problem transformed by principal component analysis. It can constitute an elastic net penalty item. The penalty characteristic of the elastic net penalty item is between the L_1 penalty item and the L_2 penalty item. The TL_1 penalty term is based on the literature [29] with a certain transformation in the form of the L_1 penalty term, and new parameters are added to the numerator and denominator to control its characteristics between the L_0 penalty term and the L_1 penalty term. However, the essence of sparse principal component analysis is still aimed at processing linear data, and the processing of nonlinear data has not yet achieved the desired effect.

In view of the situation that nonlinear data cannot be handled well, kernel principal component analysis can solve this problem well. For high-dimensional feature space, this method uses the kernel function to project sample points from the original space with a linear distribution. This method and PCA have been combined naturally. This obtains sample points that are linearly distributed in high-dimensional space and finally uses principal component analysis in this space. Through the addition of the kernel function, the principal component analysis problem of nonlinear data is successfully solved. At the same time, it avoids the difficulty of not being able to obtain an explicit projection function by constructing a kernel matrix. This successfully transformed the way of analyzing the data covariance matrix into the way of analyzing the kernel matrix.

In this section, this article will focus on giving another extension method based on principal component analysis. This method is also an improved method based on the thinking of other expansion methods of principal component analysis in this article. The sparse kernel principal component analysis based on the TL_1 penalty term is based on the thinking of sparse penalty and kernel principal component analysis. After learning sparse principal component analysis and kernel principal component analysis, it is logical to think of merging the two to get a sparse kernel principal component analysis in a high-dimensional space. In this way, the linear transformation of nonlinear data can be realized, and the sparse principal components can also be obtained in the feature space. This highlights the important variables and weakens the influence of some variables on the overall result, which has a simple and robust effect. Therefore, based on the method research on the principal component analysis, this paper proposes a sparse kernel principal component analysis.

3.1.2. Improvement Strategy. The sparse kernel principal component analysis with TL_1 penalty term combines the characteristics of the kernel principal component analysis as well as the sparse penalty, which can be implemented on nonlinearly distributed data. Its idea is similar to the sparse principal component analysis. What is considered in the sparse principal component analysis is to compare and sparse the linearly distributed data subjected to principal component analysis in the original space. When the covariance matrix is used to decompose to obtain the load, it is converted into an optimization problem, and then a penalty term is added to approach the partial load to zero. Therefore, through similar considerations, the original variables are transformed in the nuclear principal component analysis. At the same time, the eigenvalue decomposition of the sample point covariance matrix in the high-dimensional eigenspace is required. However, since the projection function cannot be obtained explicitly, it is finally transformed into the decomposition of the kernel matrix. Therefore, when the kernel matrix is eigen-decomposed, it can be transformed into an optimization problem. Then, we add the TL_1 penalty term to this optimization problem, so that the new coordinates of the sample points in the feature space are sparse.

The characteristic equation for solving the kernel principal component analysis is

$$Kv_i = f_i v_i, \tag{1}$$

where K is the kernel matrix and f_i is the eigenvalue corresponding to the eigenvector v_i.

Then, the above equation can be written as an optimized regression model:

$$\widehat{v}_i = \underset{v_i}{\arg\min} \, K - f_i v_i v_{i2}^{T2}. \tag{2}$$

Assuming $v_i = f_i v_i$ and $\alpha_i = v_i$, then equation (2) can be written as

$$(\widehat{\alpha}_i, \widehat{v}_i) = \arg\min K - v_i \alpha_{i2}^{T2}. \tag{3}$$

So far, this paper has transformed the principal component loading problem in nuclear principal component analysis into a regression optimization model. Therefore, the TL_1 penalty term can be added to the above optimization problem, so that v_i can be sparse. The model based on TL_1 penalty term sparse kernel principal component analysis is obtained as

$$\widehat{v}_i = \underset{v_i}{\arg\min} \, K - v_i \alpha_{i2}^{T2} + \lambda_1 \frac{(b+1)v_{i1}}{b + v_{i1}}, \tag{4}$$

where λ_1 and b are the parameters that need to be customized.

So far, this paper has given a sparse kernel principal component analysis model based on the TL_1 penalty term. It can be seen from this model that its idea is similar to sparse principal component analysis. The biggest difference is that sparse principal component analysis is aimed at the covariance matrix of the data. The sparse kernel principal component analysis based on the TL_1 penalty term is aimed at the mapped kernel matrix. From the perspective of the threshold iteration method, the threshold iteration algorithm for solving the sparse kernel principal component analysis based on the TL_1 penalty term is as follows:

> Step 1: normalize the input matrix to get the output matrix X. Step 2: use the kernel method to process the matrix X to obtain the kernel matrix K in the high-dimensional space. Step 3: use general principal component analysis on the matrix K to get the load matrix V. Step 4: let A' be the first k-dimensional column vector of the load matrix V; that is, $A' = V[, 1: k]$. Step 5: let the initial coefficient of the i-th principal component be x_i, given reasonable parameter values. Step 6: compare the size relationship between the parameters, and then perform the corresponding processing. Step 7: repeat Step 4 and Step 5; stop iterating when the conditions are met.

Therefore, this paper presents the basic model and algorithm of sparse kernel principal component analysis with TL_1 penalty terms. It can be seen from the above steps that the penalty functions used are all TL_1 penalty items. The process is basically similar to the solution of TL_1 penalty

term sparse principal component analysis. However, the most essential difference between the sparse kernel principal component analysis with TL_1 penalty term and the sparse principal component analysis with the TL_1 penalty term in the solution process is whether the original data has undergone a nonlinear transformation, that is to say, whether the data is mapped to the high-dimensional space. If it has undergone nonlinear processing, it is necessary to decompose and iterate the kernel matrix. Without this step, you only need to iteratively decompose the covariance matrix.

In general, an iterative algorithm is a more direct algorithm for solving optimization models. Iterative algorithms do not need to give a specific algebraic solution. Instead, it reaches a local optimal or global optimal solution by going through multiple cycles and iterations in a certain direction until it converges. However, there are many aspects that need attention in iterative algorithms. First, the selection of parameters in the iterative algorithm plays an important role and may even affect the selection of the final optimal solution. Therefore, a suitable method is needed to help the selection of parameters. Secondly, because the iterative algorithm achieves the purpose of solving by means of continuous loop iterations, so it is deeply affected by the number of iterations and accuracy. There may be a relatively large amount of calculation, especially when iterating all the elements in each vector of a matrix. Therefore, it can determine the appropriate threshold and number of iterations.

3.2. Improved BP Network

3.2.1. Improved Particle Swarm Optimization. After using the proposed principal component analysis method to reduce the dimensionality of the features, this paper uses the BP network to model these features to predict the popularity of news public opinion.

Prediction accuracy is low due to random selection of initial threshold and weight parameters and the BP network's tendency to slip into the local minimum state. This results in poor performance. BP neural network model predictions can be enhanced by using an upgraded particle swarm technique to optimize the starting parameters. Model predictions can be made using a more accurate BP model that has been optimized for use in this study.

Particle swarm optimization (PSO) is an algorithm based on swarms that is extremely clever. The algorithm is easy to develop, is precise, and quickly converges to the best solution. It can effectively optimize various functions and is suitable for application in multiobjective constrained optimization problems. And it has better efficiency and practicability in solving various problems. At present, it has achieved good application results in constrained optimization, function optimization, various engineering design problems, and other fields.

Based on population dynamics, particle swarm optimization (PSO) is a population-based approach for solving optimization problems. One of its most fundamental aspects involves the creation of an initial group of particles in space using a randomized search process. Iteration is used to find

the best solution for these particles, which begins with a random solution. There are three ways to express the attributes of a particle in an optimization problem: position, speed, and the fitness value. The fitness function can be used to calculate the particle's fitness value, which can be used to determine the particle's advantages and disadvantages. During each repetition, the particle updates itself by observing two points of departure. The individual extreme value is the one being sought by the particle, whereas the global extreme value is the one being sought by the entire group at the same time. Particles modify their speed and position by following these two extreme values, which each particle follows.

The basic principle of PSO mainly includes the initialization of particle swarm, the calculation for particle fitness, the update of particle speed and position, and the determination of termination conditions. The specific steps are as follows:

> Step 1: initialize the particle swarm. Step 2: calculate the objective function value, namely, fitness. Step 3: calculate fitness after the location update. The fitness is analyzed and compared with its individual extreme value Pbest. If fitness > Pbest, replace Pest with fitness. Step 4: compare fitness with the global extreme value Gbest. If fitness > Gbest, replace Gbest with fitness. Step 5: update the speed and position. Step 6: judgment of termination conditions. If the algorithm progresses until the end condition is met, the iterative calculation process is stopped, and the result is output at this time. Otherwise, the algorithm returns to Step 2.

A particle swarm algorithm can better solve combinatorial optimization problems. But this algorithm also has some shortcomings. Algorithm search accuracy is not very great and it is easy to get caught in the trap of local minimum solutions when using it. A modified particle swarm method using inertia weight has been developed to address these problems; in this way, the algorithm's search performance can be improved and it can avoid getting stuck in a local optimal scenario.

Among the various parameters of PSO, inertia weight w is an important parameter. This parameter affects the local optimal ability as well as the global optimal ability, and its value can reflect the speed of the particle's flight. If, in the algorithm parameter setting, the value for inertia weight w increases, the flying speed will increase, which can enhance the global search capability. On the contrary, if the value of w is reduced, the flying speed of the particles will slow down, which is beneficial to improve the local searchability. Therefore, it is necessary to select an accurate inertia weight w, so that the particles have balanced searchability, and effectively improve the exploration ability of the algorithm.

At present, many scholars have proposed improved methods for inertial weights, such as random weighting algorithms and fuzzy inertial weighting algorithms. One of the more common algorithms uses the method of linearly decreasing weights. Aiming at the characteristic that the PSO algorithm is easy to mature prematurely, literature [30] proposed a linear decreasing weight method. By introducing

a linear transformation weight, the inertia weight decreases linearly from large to small. The formula for change with the number of iterations is as follows:

$$w = \frac{w_{\max} - (w_{\max} - w_{\min})t}{t_{\max}}, \quad (5)$$

where $w_{\max}$ is the maximum inertia weight, $w_{\min}$ is the minimum inertia weight, t represents the current iteration, and $t_{\max}$ is the maximum iteration.

Through analysis, although the introduction of linear inertia weight can make the PSO algorithm adjust the particle's optimization ability, this algorithm has some shortcomings. First of all, in the early stage of operation, due to the large value of the inertial weight w parameter, it is conducive to the global search of particles. If the optimal point is detected in the early stage of the algorithm, it is hoped that the algorithm will quickly converge to its optimal point. However, the linear decreasing characteristic of the inertia weight w slows down the convergence speed of the algorithm. Secondly, in the late running of the algorithm, as w decreases, the global search ability gradually decreases, which is easy to fall into the local optimal situation at this time.

Aiming at the shortcomings, an improved method of nonlinear decrease of inertia weight is proposed on the basis of the algorithm to improve the shortcomings. This algorithm is described as

$$w = w_{\max} - (w_{\max} - w_{\min}) \times \tan\left(\frac{\pi t}{4t_{\max}}\right). \quad (6)$$

In the iterative process of the algorithm, when t is small, the global search capability can be guaranteed. Then, with the increase of t, local search ability can be guaranteed, so the algorithm has strong balanced searchability. Through the above-mentioned improved method, the nonlinear dynamic inertia weights better realize the balance between the global search and the local search. This improves the performance in convergence speed and global optimization.

3.2.2. Improvement Strategy for BP.

BP network is an algorithm based on gradient descent, as shown in Figure 1. As a result, the training effect of the network model will be negatively impacted by its inability to search globally. This cannot provide a more effective training method. Global optimization and simultaneous search are two features of the particle swarm algorithm. This has the potential to enhance global convergence. Neural networks can be improved by using this technique in the neural network algorithm. The particle swarm algorithm's global optimization ability is merged with the neural network's advantages by comparing their properties. It is possible to improve the learning and training speed of the network by establishing a BP neural network model using the particle swarm algorithm and optimizing the initial weights and threshold parameters of the BP network.

PSO may be used to optimize the initial parameters of a BP neural network model by integrating the particle swarm
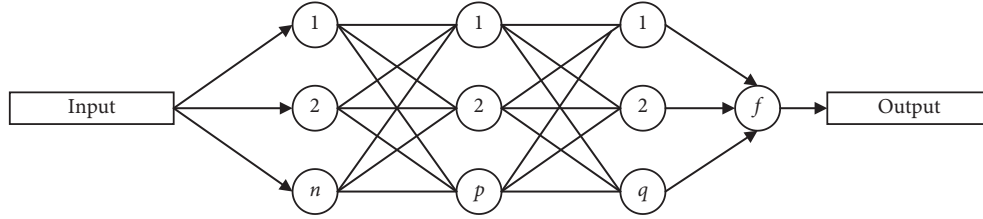
Figure 1: Structure of the BP network.

4. Experiment and Discussion

method into the model. Iterative particle optimization can be used to find the best parameters for the BP network. When it comes to the PSO algorithm, the BP neural network weights and thresholds are the flying particles of dimension two. In order to find the most optimal particles, you must first select the best initial parameters for the BP network. Additionally, iterative particles constantly adjust their velocity, position, and initial characteristics in space. Searching for the ideal particle is done in accordance with the function's optimization aim so that the particle's optimal parameter can be determined. The fitness value of a particle is the difference between the projected output of the BP model and the expected output, which is the optimization objective function. The fitness function is

$$f = \frac{1}{n} \sum_{i=1}^{n} |y_i - p_i|, \tag{7}$$

where n is the number of samples, y_i is the actual output, and p_i is the predicted output.

Through the continuous update for the position, velocity, weight, and threshold, we determine the parameter value when the particle obtains the optimal fitness, assign the parameter value, and use the parameter optimized model for learning and training.

The principle of using the PSO algorithm to optimize the BP network algorithm mainly includes three parts: determining BP network structure, the PSO algorithm optimizing BP network parameters, and BP network training and prediction. The specific steps are as follows. (1) First, initialize the parameters for the model, including the number of network layers and learning rate of BP network, the number of populations of particle swarm algorithm, inertia weight, and the number of iterations. (2) Calculate and compare the fitness value. The output error of BP network model training is used as the fitness objective function to obtain the fitness value. (3) According to the fitness function, the improved particle swarm algorithm is used to iteratively search the space and update speed and position. (4) Determination of the termination condition of particle population update: when the algorithm reaches the maximum iterations, the loop iterative calculation process is stopped, the optimal fitness value is determined, and the optimal initial parameters are obtained. (5) Assign the optimized parameters to the BP network as initial weight and threshold parameters of the BP model. Combined with the sample data of news and public opinion, the optimized BP model is used to train until the entire model meets the convergence accuracy. (6) Use the trained BP model to predict the popularity of news public opinion.

4. Experiment and Discussion

4.1. Dataset. This article self-made two data sets DS1 and DS2. DS1 comes from online news published by Mashable Publishing House, and a total of 1389 news data are selected. DS2 comes from news published by Facebook, and a total of 1473 news data are selected. The specific distribution of the training set and test set is shown in Table 1. For each piece of news data, a 60-dimensional feature is extracted, and the dimensionality is reduced to a 20-dimensional feature vector through an improved principal component analysis algorithm. Each piece of news is divided into two categories: popular and unpopular. In this work, precision, recall, and *F*1 score are utilized to evaluate the performance of prediction.

4.2. Evaluation of Model Convergence. In the BP network, whether the model converges is an important indicator for evaluating network performance. If the model fails to converge, subsequent predictions are meaningless. Therefore, this article first compares the training loss of the proposed network on two data sets. Experimental results are illustrated in Figure 2.

As the training progresses, the loss of the network gradually decreases. And at the 40th epoch, the loss is basically not decreasing, indicating that the network has reached a state of convergence. Besides the training loss, this work also compares the test performance on two datasets. Experimental results are illustrated in Figure 3.

Similarly, when the training iteration is 40 epochs, the performance of the network on the test set tends to converge, which can obtain 0.91 precision, 0.82 recall, and 0.86 *F*1 on DS1, 0.93 precision, 0.85 recall, and 0.88 *F*1 on DS2. In summary, the designed network can finally converge and make stable and efficient predictions.

4.3. Evaluation on Improved PCA. In this work, an improved PCA is proposed to reduce the dimensionality of features. To prove that this method can effectively improve the performance, the performance when using the improved PCA (IPCA) is compared with that of the improved BP (IBP) network alone. The experimental results are illustrated in Figure 4.

Combining the IPCA method with the BP network can effectively improve the performance. On the DS1 dataset, compared with not using the IPCA method, the performance improvement obtained is 6%, 5%, and 5% on three indexes. On the DS2 dataset, compared with not using the

TABLE 1: The division of training set and test set.

Dataset	Training set	Test set
DS1	1041	348
DS2	1104	369
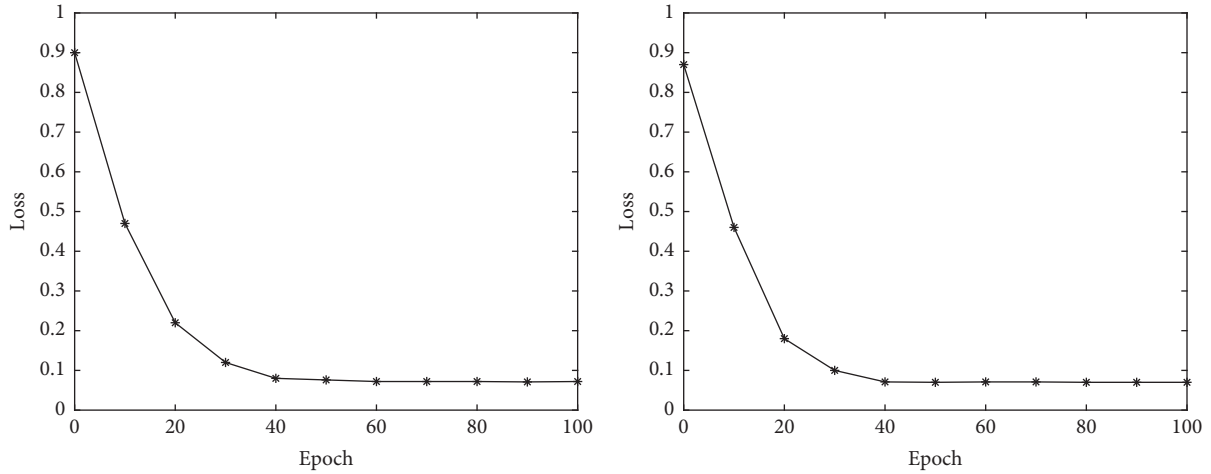
FIGURE 2: The training loss on DS1 and DS2.
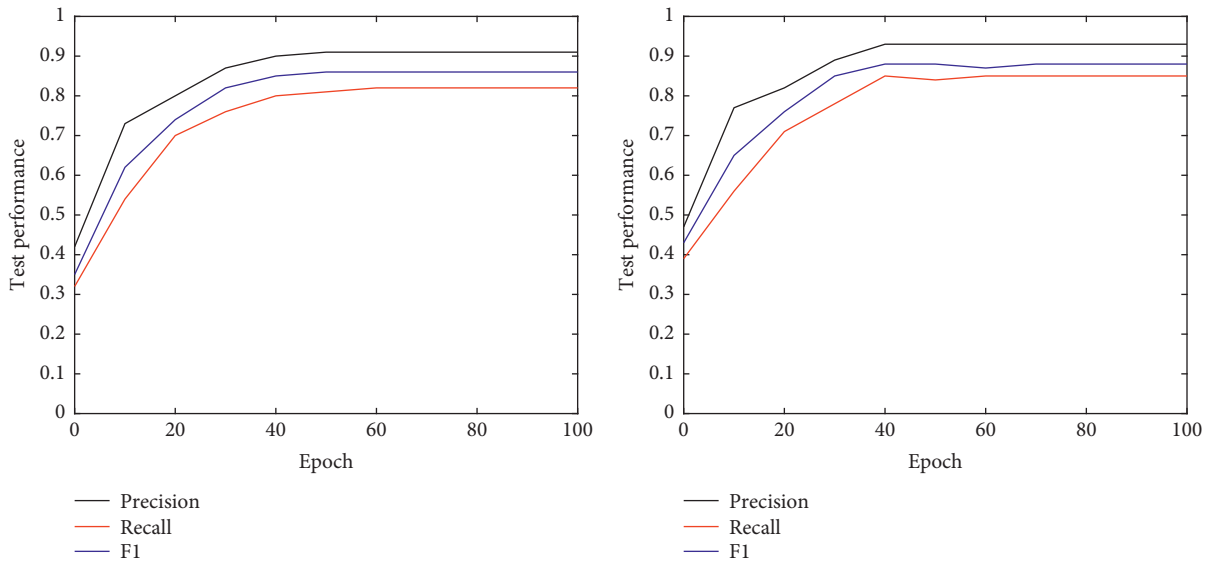
FIGURE 3: The test performance on DS1 and DS2.

IPCA method, the performance improvement obtained is 7%, 6%, and 5% on three indexes. To further illustrate the superiority of the IPCA method, it is compared with the traditional PCA. Experimental results are illustrated in Table 2.

Obviously, compared with the PCA algorithm, the IPCA method can further improve the performance of the BP network. This proves the reliability and effectiveness of the IPCA method.

4.4. Evaluation of Improved PSO.

In this paper, an improved PSO is proposed to optimize the BP network. To prove that this method can effectively improve the performance of the BP network, the performance when using the improved PSO (IPSO) is compared with that of the BP network alone. The experimental results are illustrated in Figure 5.

It can be seen that combining the IPSO method with the BP network can improve performance. On the DS1 dataset, compared with not using the BP method, the performance
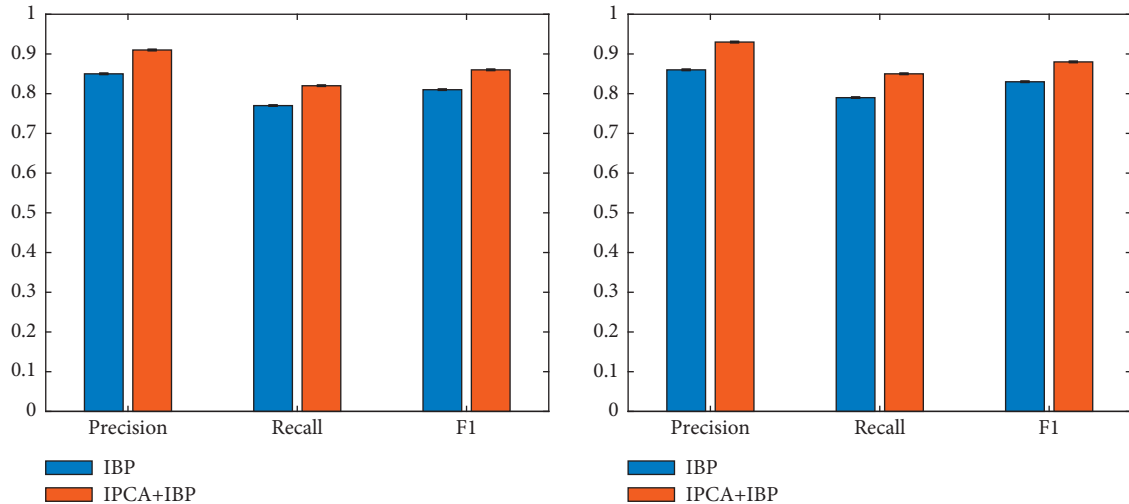
FIGURE 4: Evaluation of improved PCA.

TABLE 2: Comparison of IPCA and PCA.

Method	DS1			DS2		
	Precision	Recall	$F1$	Precision	Recall	$F1$
PCA + IBP	0.89	0.80	0.83	0.90	0.82	0.84
PCA + IBP	0.91	0.82	0.86	0.93	0.85	0.88
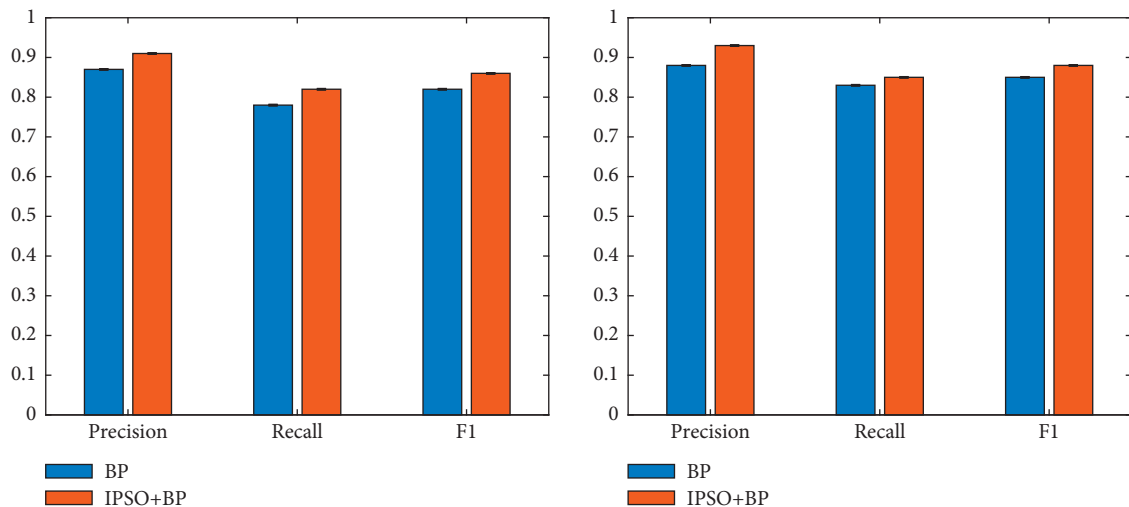
FIGURE 5: Evaluation of improved PSO.

improvement obtained is 4%, 4%, and 4% on three indexes. On the DS2 dataset, compared with not using the IPSO method, the performance improvement obtained is 5%, 2%, and 3% on three indexes. To further illustrate the superiority of the IPSO method, it is compared with the traditional PSO method. Experimental results are illustrated in Table 3.

Obviously, compared with PSO, the IPSO method can further improve the performance of the BP network. This proves the reliability and effectiveness of the IPSO method in this paper.

TABLE 3: Comparison of IPCA and PCA.

Method	DS1			DS2		
	Precision	Recall	$F1$	Precision	Recall	$F1$
PSO + BP	0.88	0.80	0.84	0.91	0.84	0.85
IPSO + BP	0.91	0.82	0.86	0.93	0.85	0.88

4.5. Comparison to Other Methods. To illustrate the effectiveness of our strategy, this article compares our method with other methods. Experimental results are illustrated in Table 4.

TABLE 4: Comparison to other methods.

Method	DS1			DS2		
	Precision	Recall	F1	Precision	Recall	F1
DT	0.75	0.67	0.71	0.79	0.69	0.73
SVM	0.81	0.72	0.75	0.81	0.78	0.75
BP	0.85	0.77	0.79	0.86	0.83	0.82
Ours	0.91	0.82	0.86	0.93	0.85	0.88

Obviously, compared with other methods, our method can obtain the best performance improvement. This demonstrates the correctness of the method in this article.

5. Conclusion

With the development of the Internet, news and public opinion have exploded. It is a very important subject to extract and study the characteristic attributes that determine the popularity of news public opinion and establish a network model to predict the popularity of news public opinion. The main work of this paper has the following two contents. The first is to realize the sampling of news and the extraction of attribute features. Then, considering the nonlinear relationship between features, an improved principal component analysis method is proposed to analyze the correlation of features. In this way, important features can be selected from many irrelevant features, effectively reducing the original high-dimensional features. Second, neural networks are examined for their ability to anticipate public opinion on news stories. By using BP, we can mine data more effectively. The BP neural network, on the other hand, has a number of limitations and uses an improved approach for optimizing the initial parameters of the BP neural network, which successfully compensates for the BP neural network's flaws. Finally, a prediction model for the popularity of news public opinion is built using the BP network with optimal parameters. The new neural network model described in this paper can accurately predict the popularity of news public opinion, as demonstrated by experimental findings.

Data Availability

The datasets used are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] T. C. Serm, P. Blanchfield, and K. Su, "Mobile newspaper development framework: guidelines for newspaper companies for creating usable mobile news portals," in *Proceedings of the 2006 International Conference on Computing & Informatics (ICOCI 2006)*, pp. 1–8, IEEE, Kuala Lumpur, Malaysia, June 2006.

[2] W. Wang, "News development in the 5G network era based on machine learning and FPGA," *Microprocessors and Microsystems*, Article ID 103391, 2020.

[3] B. H. Shah, M. Adnan, and I. Sultana, "Development news in print media of Pakistan: the case of Urdu and English press," *SJESR*, vol. 3, no. 2, pp. 244–251, 2020.

[4] C. A. Watson, "Information literacy in a fake/false news world: an overview of the characteristics of fake news and its historical development," *International Journal of Legal Information*, vol. 46, no. 2, pp. 93–96, 2018.

[5] B. P. Putra, K. Ansari, and I. Pramuniati, "The development of news text-based problems teaching materials in 12th grade students of vocational high school at TR sinar husni of 2019/2020," *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, vol. 3, no. 1, pp. 293–305, 2020.

[6] J. M. Cleaton, C. Viboud, L. Simonsen, A. M. Hurtado, and G. Chowell, "Characterizing ebola transmission patterns based on internet news reports," *Clinical Infectious Diseases*, vol. 62, no. 1, pp. 24–31, 2016.

[7] Y. Zhang, W. Song, D. Shen, and W. Zhang, "Market reaction to internet news: information diffusion and price pressure," *Economic Modelling*, vol. 56, pp. 43–49, 2016.

[8] J. W. Ayers, B. M. Althouse, M. Dredze, E. C. Leas, and S. M. Noar, "News and internet searches about human immunodeficiency virus after charlie sheen's disclosure," *JAMA Internal Medicine*, vol. 176, no. 4, p. 552, 2016.

[9] J. Pei, K. Zhong, and J. Li, "ECNN: evaluating a cluster-neural network model for city innovation capability," *Neural Computing & Applications*, pp. 1–13, 2021.

[10] A. C. Micu and I. Pentina, "Examining search as opposed to experience goods when investigating synergies of internet news articles and banner ads," *Internet Research*, vol. 25, no. 3, pp. 378–398, 2015.

[11] Y. Huang, Z. Yu, J. Guo, Z. Yu, and Y. Xian, "Legal public opinion news abstractive summarization by incorporating topic information," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 9, pp. 2039–2050, 2020.

[12] C. Yasai, "The formation of Chinese online public opinion on the 7.23 wenzhou train crash: through the analysis of sina news site, sina weibo, and tianya BBS," *Journal of Mass Communication Studies*, vol. 86, pp. 123–142, 2015.

[13] M. Yousaf, W. Li, and Z. Yousuf, "The news media exposure and polarization of public opinion: the limits of social functions of the agenda-setting role of mass media," *Global Media Journal*, vol. 13, no. 1, pp. 36–51, 2020.

[14] O. C. Dewi and R. M. Amalia, "Construction of public opinion in online news on the issue of nurses' expulsion from boarding house: a critical discourse analysis," *Teknosastik*, vol. 19, no. 1, p. 48, 2021.

[15] I.-H. Hong, "Analysis of public opinion on sports by the 4th industrial revolution using newspaper big data," *Korean Journal of Sport Management*, vol. 25, no. 3, pp. 1–9, 2020.

[16] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proceedings of the ACM SIGKDD International Conference on Knowledge*

Discovery and Data Mining, pp. 497–506, Paris, France, June 2009.

[17] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 US election: divided they blog," in *Proceedings of the International Workshop on Link Discovery*, pp. 36–43, Chicago, IL, USA, August 2005.

[18] M. Tsagkias, W. Weerkamp, and M. De Rijke, "Predicting the volume of comments on online news stories," in *Proceedings of the ACM Conference on Information and Knowledge Management*, pp. 1765–1768, Beijing, China, 2009.

[19] A. Kaltenbrunner, V. Gómez, and V. López, "Description and prediction of slashdot activity," in *Proceedings of the Latin American Web Conference*, pp. 57–66, Santiago, Chile, 2007.

[20] G. Szabo and B. A. Huberman, "Predicting the popularity of online content," *Communications of the ACM*, vol. 53, no. 8, pp. 80–88, 2010.

[21] M. Tsagkias, M. De Rijke, and W. Weerkamp, "Linking online news and social media," in *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 565–574, Hong Kong, China, 2011.

[22] K. Lerman and R. Ghosh, "Information contagion: an empirical study of the spread of news on digg and twitter social networks," *Computer ENCE*, vol. 52, pp. 166–176, 2010.

[23] J. G. Lee, S. Moon, and K. Salamatian, "An approach to model and predict the popularity of online contents with explanatory factors," in *Proceedings of the International Conference on Web Intelligence and Intelligent Agent Technology*, pp. 623–630, Toronto, Canada, August 2010.

[24] R. Bandari, S. Asur, and B. Huberman, "The pulse of news in social media: forecasting popularity," in *Proceedings of the AAAI Conference on Web and Social Media*, pp. 12–17, Dublin, Ireland, 2012.

[25] A. Tatar, P. Antoniadis, M. D. d. Amorim, and S. Fdida, "From popularity prediction to ranking online news," *Social Network Analysis and Mining*, vol. 4, no. 1, p. 174, 2014.

[26] M. Tsagkias, W. Weerkamp, and M. De Rijke, "News comments:exploring, modeling, and online prediction," *Lecture Notes in Computer Science*, pp. 191–203, 2010.

[27] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, UK, 2003.

[28] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[29] M. Nikolova, "Local strong homogeneity of a regularized estimator," *SIAM Journal on Applied Mathematics*, vol. 61, no. 2, pp. 633–658, 2000.

[30] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the IEEE International Conference on Evolutionary Computation*, pp. 69–73, Anchorage, AK, USA, 1998.

Research Article

Design of Building Environment Detection System for Architectures Based on Internet of Things

Dongfang Zhang,[1] **Hui Ji,**[2] **Zhennan Li,**[1] **and Hui Ge** [1]

[1]*School of Planning and Design, Xinyang Agriculture and Forestry University, Xinyang, Henan 464000, China*
[2]*School of School of Architecture and Urban Planning, Guangdong University of Technology, Guangzhou,*
 Guangdong 510000, China

Correspondence should be addressed to Hui Ge; 2002230021@xyafu.edu.cn

In the process of urban building design, a new integrated system for monitoring the environment is developed and designed by using embedded development technology and sensor technology. The system uses a wireless sensor network environment monitoring system IoT platform with embedded internal processors. Analyze and design the system as a whole, including the construction of the basic platform of the system, the design of the internal plates and circuits of the system, the monitoring design of the input node, and the monthly design of the output interface calculation. Finally, a physical model is built, and data measurement and analysis are carried out under different conditions, and the evaluation and advantage analysis of the system's operating status are given. The system can carry out all-round, multilevel, and three-dimensional real-time monitoring of the construction site environment, including dust, *PM2.5*, temperature, humidity, wind speed, carbon dioxide, and other indicators in the construction site environment. In addition, the system can upload various monitoring data to the detection system through the internal network. The system has the functions of monitoring, alarming, recording, querying, and counting of the target monitoring station and can also be linked with the environmental control device. The construction site staff can conduct real-time supervision through the mobile terminal and computer terminal management platform. In addition, it can also meet the role of real-time remote monitoring and online guidance and regulation. It has reference value for the safety and management of the actual operation process of the project.

1. Introduction

The construction industry has always been a fast-growing industry. With the emergence of more and more modern high-tech, building environment monitoring system has also become a hot research topic in the field of construction. Before this, people could only passively adapt to the external environment using the materials in the living environment. However, due to the gradual development of science and technology, people have gradually become active in transforming the living environment and have taken active measures to improve the environment of various buildings in terms of building site selection, planning and design, building envelopes, building equipment, and building materials [1]. Most of the office buildings in our country have

serious energy consumption, the unit energy consumption is gradually increasing, and its energy expenditure is also increasing. This requires research on a method to reduce the energy consumption of buildings. Although many scholars have studied building energy consumption monitoring systems, there are many energy consumption monitoring points in buildings and complex wiring. In this way, the monitoring effect is not good, and the energy consumption of the building cannot be effectively reduced. Based on this problem, on the basis of the building site environmental monitoring software we mentioned above, we design and regulate a building comprehensive environmental energy consumption monitoring and management system based on the Internet of Things technology to solve the above problems [2]. In the traditional building environment

monitoring field, the system usually uses a distributed monitoring system composed of an upper computer and a lower computer to monitor the building environment, and the lower computer and various sensors use network communication. This type of monitoring system has the characteristics of flexible structure, easy expansion and upgrading, and low cost [3]. However, with the continuous development of building intelligence, people's requirements for building environment monitoring are also increasing, which usually requires long-distance remote monitoring and big data analysis and decision-making. Therefore, if the traditional environment monitoring system architecture is still adopted, the disadvantages such as heavy wiring workload, troublesome data processing, and unsuitability for long-distance transmission are prominent. With the continuous development of Internet of Things technology and wireless transmission technology, it is possible to design a wireless building environment monitoring system with little or no wiring workload [4].

The role of the Internet of Things has been used in industry, medical care, agriculture, logistics, transportation, power grid, environmental protection, fire protection, furniture, and other aspects, which closely connected with life. With the continuous development and maturity of the technology, the Internet of Things will bring revolutionary changes to life in the future and help human beings to step into the era of wisdom faster and more stably. With the development and improvement of information technology, sensor technology, and Internet of Things technology, it is possible to build an integrated system suitable for long-term monitoring of building environment. In the actual operation state of the building, the performance data such as the energy consumption of the built environment and the indoor environment and the user satisfaction evaluation are collected, and the actual operation effect of the green building is solved through analysis and identification. In this way, it is beneficial to move towards refined green building design and operation optimization. Therefore, it is very necessary to conduct long-term monitoring of the indoor and outdoor environment of the building, fully understand the mechanism, and then analyze the monitoring data and use relevant means to optimize the building environment [5]. IoT technology will reduce site injuries and make building construction more efficient. As the types of monitoring equipment become more and more complex, fieldbus technology can manage different types of monitoring equipment in a unified manner. At present, RS-485 technology is mostly used to monitor the building environment and meteorological conditions, in order to stably and efficiently monitor the environment of the solar greenhouse in the northern cold region. In addition, RS-485 can be used to establish a wired data transmission network for the acquisition node and data transmission between the acquisition node and the sensor group through the Modbus communication protocol and then use the 433 MHz wireless data acquisition device to connect with remote servers and clients [6–9].

In this design system, another latest invention technology is used: wireless sensor network technology. Wireless sensor network technology has the advantages of convenience, flexibility, and wide application. Therefore, in recent years, many scholars have built relatively complete wireless sensor networks for environmental monitoring. With the help of wireless sensor network technology, that is, Arduino devices and VOLTTRON software platform to integrate ZigBee protocol sensors, a low-cost building performance monitoring system toolbox can be established. In addition, with the emergence of ZigBee protocol and Wi-Fi to form wireless sensor network technology, combined with the back-end monitoring information of IoT components, a monitoring system can be formed as an effective service platform [10]. From Figure 1, we can see that the types of building environment monitoring sensors mainly include temperature sensors, humidity sensors, formaldehyde sensors, haze sensors, VOC sensors, noise sensors, etc. Various types of sensors are usually composed of sensing modules, information processing modules, wireless communication modules, and energy supply modules [11]. The building energy consumption and environmental monitoring system based on the Internet of Things technology monitors the five energy systems of electricity, water, gas, heat consumption, and cooling consumption in the building energy consumption and analyzes the energy consumption data in combination with the environmental monitoring system inside the building. In the realization technology, the field bus instrument is used to realize the data acquisition of the 5 types of energy systems. Environmental parameter monitoring is implemented using ubiquitous wireless sensor network-based detectors. The construction industry is bringing real-time information into centuries-old processes. The devices and sensors of Internet of Things (IoT) are gathering workplace data in ways that are more affordable, efficient, and effective than previously thought.

Adopt industrial database and configuration software to realize the summarization and analysis of the collected data [12]. This system is based on computer, communication equipment, measurement, and control unit as basic tools and provides a basic platform for real-time data acquisition, switch status monitoring, and remote management and control of large public buildings. It can form any complex monitoring system with detection and control equipment [13]. The system mainly adopts a layered distributed computer network structure, which is generally divided into three layers: the station control management layer, the network communication layer, and the field device layer [14]. The field device layer is the data acquisition terminal, which is mainly composed of intelligent instruments. From Figure 2, we can see that the distributed I/O controller with high reliability and fieldbus connection is used to form the data acquisition terminal, and the stored building environment monitoring data is uploaded to the data center. Measuring instruments are responsible for the most basic data collection tasks, and the effective data monitored by them must be complete, accurate, and transmitted to the data center in real time.

2. The Composition and Principle of Radio Frequency Identification Technology

The composition of the Radio Frequency Identification (RFID) system includes four aspects:
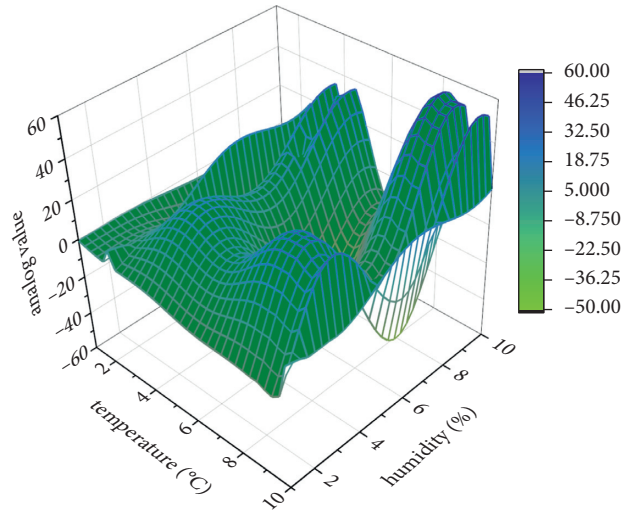
FIGURE 1: Hoflander's attitude change-persuasion model.
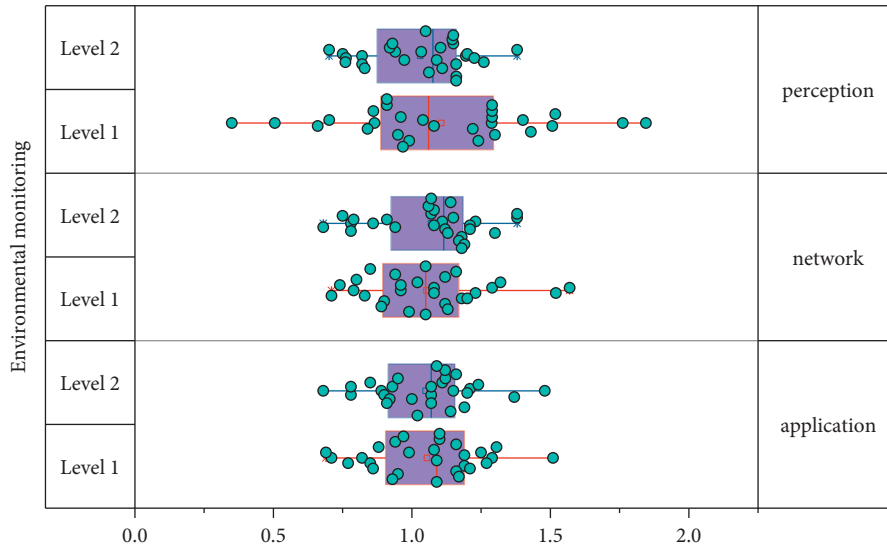
FIGURE 2: Calculation results of three levels of environmental monitoring.

(1) Reader: It sends radio frequency signals to identify the electronic tag. After the identification is successful, it receives the information transmitted by the electronic tag and transmits it to the computer system for data analysis [15].

(2) Electronic tag: It stores the relevant data information of the target to be identified. According to the energy source of the electronic tag, it can be divided into active tags, passive tags, and semiactive tags.

(3) Antenna: including transmitting antenna and receiving antenna, which are used for energy and information transmission between readers and electronic tags and are an important part of RFID system. The quality of antenna performance directly affects the efficiency of energy transmission.

(4) Computer system: data integration, processing, and analysis of the information received by the reader [16].

$$As = A|C|^{-\partial}. \tag{1}$$

In the formula, As is the sensor resistance, A is a constant, and C is the gas concentration.

The RFID system principle is composed of the following requirements. Firstly, a section of inquiry command is encoded, and the signal is modulated, and the radio frequency signal is sent to the electronic tag through the transmitting antenna, and finally the information stored by the tag is received by the receiving antenna [17]. This process enables the transfer of energy while simultaneously decoding, demodulating, and rectifying the received signal. From Figure 3, we can see that passive electronic tags are excited by energy, encode and modulate the information carried by themselves, and resend it to the reader, and the reader receives the information. Finally, the secondary
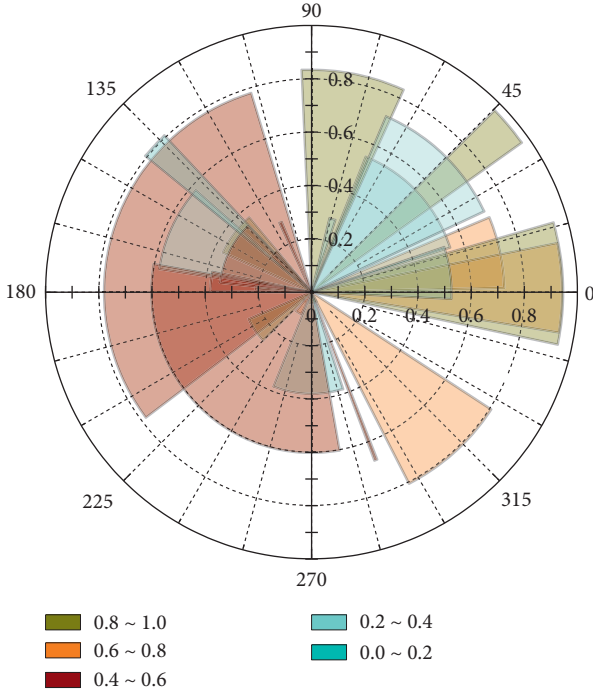
Figure 3: SVM training set classification accuracy.

information is submitted to the computer system for background processing, so as to complete the automatic identification work, and both modulation and demodulation are completed in the *RF* modem.

The transmission of signals between readers and electronic tags in an *RFID* system takes two types. The first is the inductive coupling model. This model is suitable for *RF* signals in the middle and low frequency bands. Similar to the transformer model, the magnetic induction loop of the electronic tag will be affected by the magnetic field change of the reader coil, which will generate an induced electromotive force and supply energy for the electronic tag, as shown in Figure 3, where the dotted box represents the magnetic field. The second is the electromagnetic backscattering model. This model is suitable for *RF* signals in the high frequency and microwave frequency bands. Compared with the previous model, due to the short wavelength of the radio frequency signal sent by the magic reader, after radiating to the electronic tag, part of it will carry the tag information and be reflected back to the reader. It is interesting that this model works similarly to the detection radar model [18].

$$f(x) = \text{sign}\left(\sum_{t=1}^{m} a y_t k(x_t, x_j) + b\right). \tag{2}$$

In the above formula, since the environmental data is nonlinear, the samples of the original input space should be mapped to the high-dimensional feature space, and then the optimal classification hyperplane should be constructed.

Compared with the traditional automatic identification technology, *RFID* technology has broad development prospects. People's research on sensors pays more attention to technology fusion. For example, by using the fusion of

RFID technology and sensor technology, the sensor and the monitoring terminal can communicate and transmit information through radio frequency signals. From Figure 4, the data on the way accord with the relation of exponential function. From the distribution of fitting curve and actual data points, the mathematical model used in this study can accurately predict the trend of data development. This model can well achieve real-time monitoring and early warning of various indoor and outdoor toxic and harmful gas concentrations. In addition, it can effectively identify the concentration changes of some toxic and harmful gases over a period of time [19]. Based on the principle of this design procedure, the optimal allocation of energy can be effectively realized. In addition, the wireless networking mode and modular structure in the design are simple, the networking is convenient, and the construction and reconstruction are convenient. The visual interface provides a more convenient operation management mode for construction personnel, building management, and users and is easy to promote. This shows that the software can be used as an effective method for building green and energy-saving buildings [20].

3. Design of Automated Supervision Platform

3.1. Key Technologies of Automated Supervision Platform. First, establish the application and demonstration of environmental automation supervision on construction sites covering urban residential areas, and apply Internet of Things technology and cloud computing technology to the data center. From Figure 5, we can see that, after data analysis, statistics, spectrum analysis, storage, recording processing, and meteorological parameters, it is transmitted to the construction site project supervision command center, and various monitoring data are displayed in the real time at the construction site for the supervision of surrounding residents.

Second, establish a supervision cloud platform for environmental protection functional departments. The cloud platform mainly includes data communication supervision system, core management system, and data exchange platform. The core management system is mainly composed of a data display platform, a supervision platform, a public service platform for information release, and a standard data access platform. It is a bridge connecting front-end smart devices and data processing centers. The data processing center is mainly a bridge between the front-end intelligent monitoring sample collection equipment and the data processing center. The data center mainly has the management and control of the front-end intelligent monitoring sample collection equipment. The core management system has the function of virtualized management of all collected monitoring data. The data exchange platform saves the data to the corresponding data processing center according to the requirements of the cloud computing data storage platform.

Third, establish a unified data access standard. For example, the environmental information monitoring data and statistical information of construction companies have spatial attributes in addition to time and dynamics. It is suitable for expression using geographic information
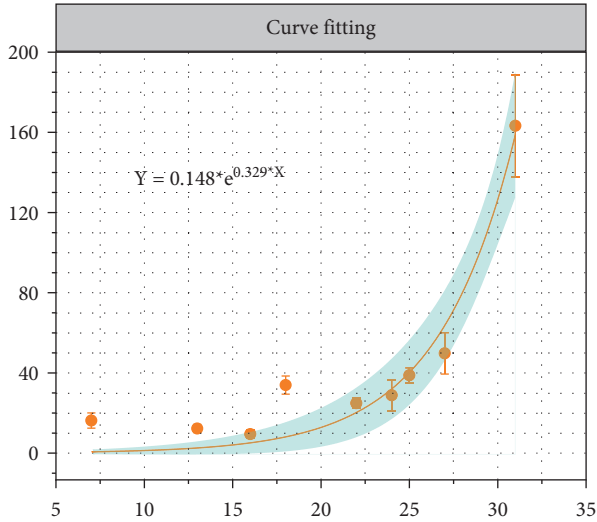
FIGURE 4: Simulated data fitting curve.
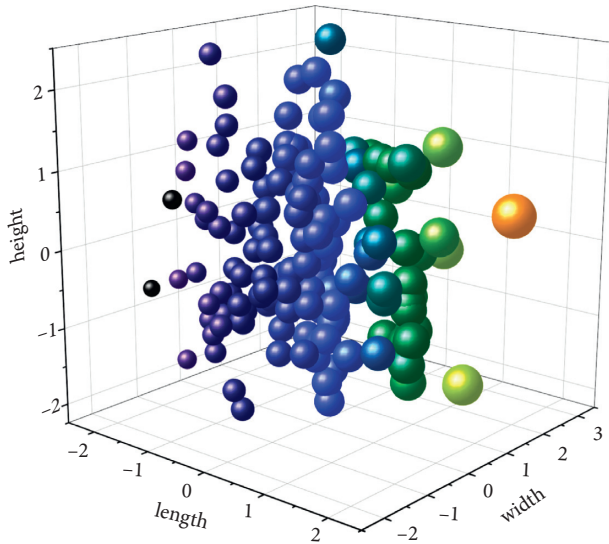
FIGURE 5: Motion recording of gas molecules in the environment.

system. This system has a powerful data interface. The *GIS* system of the management department command center can integrate the construction site environmental data through the unified platform data access interface and dynamically and intuitively display the data of each environmental monitoring point.

3.2. Environmental Quality Detection and Feedback Circuit Design

3.2.1. Sensor Module Circuit Design.
The sensor used in the environmental quality detection part is composed of *GY-MCU680V1* sensor, *YW-51 GJ* sensor, *MQ-2* sensor, and *MQ-7* sensor. *GY-MCU680V1* indoor environmental quality sensor is a multisensor which is responsible for detecting air index, temperature, humidity, and air pressure. It integrates the BME680 four-in-one environmental sensor and

processes the detected data through its own processor and outputs it through the serial port. *YW-51GJ* dust sensor is the sensor responsible for detecting *PM2.5* parameters. From Figure 6, we can see that the sensor uses the principle of scattering, by detecting the voltage value of the receiving element as a reference for the dust concentration. The sensor is integrated with an *ADC* converter and a processor, and the processor outputs the internal *ADC* conversion result through a serial port. *ZE08-CH2O* formaldehyde sensor is a sensor responsible for detecting formaldehyde concentration parameters. The sensor is responsible for measuring the formaldehyde concentration in the air and outputting it through the voltage value. The *MQ-2* and *MQ-7* sensors are hazardous gas sensors, which change the resistance of their gas-sensing resistors by reacting with the corresponding sensitive gases in the air, thereby producing voltage changes. When the voltage exceeds the limit, it will send out the corresponding high- and low-level signal through the *I/O* port to trigger the alarm.

3.2.2. Display Module Circuit Design.
The display module of the design system uses the 3.5-inch *TFT LCD* screen of Antike Technology which is matched with the core board. At the same time, the matching controller adopts *ILI9488* as the controller and uses the *FPC* cable to connect with the core board.

$$PPD = 100 - 95 \times e^{-0.03353 \times PMV^4 - 02179 \times PMV^2}. \tag{3}$$

This formula expresses the calculation formula of environmental thermal comfort evaluation.

3.2.3. Circuit Design of Alarm Module.
The alarm module is *TELESKY* 3.3~5V low-level drive active buzzer, of which *GND* and *VCC* are connected to the core board power supply circuit (3.3 V) to supply power to the module.

$$APMV = \frac{pmv}{(1 + \lambda \times pmv)}. \tag{4}$$

Among them, λ is the adaptive coefficient, which is -0.3 in the construction site.

3.2.4. Module Circuit Design.
Wi-Fi module uses Shenzhen Zhida *ESP8266* Wi-Fi module, which uses *ESP8266* processor to handle Wi-Fi corresponding operations, and uses serial port to control the processor. The *ZigBee* module uses the Deep Link Innovation *DL-LN33* ZigBee module, which uses the *CC2530* processor to handle the corresponding operations of the *ZigBee*, and uses the serial port to control the processor.

$$y = \frac{2(X_t - X_{\min})}{X_{\max} - X_{\min}} - 1. \tag{5}$$

In the formula, x represents the collected air quality data, and y represents the data after normalization and transformation, which can reduce the deviation between the result and the actual value.
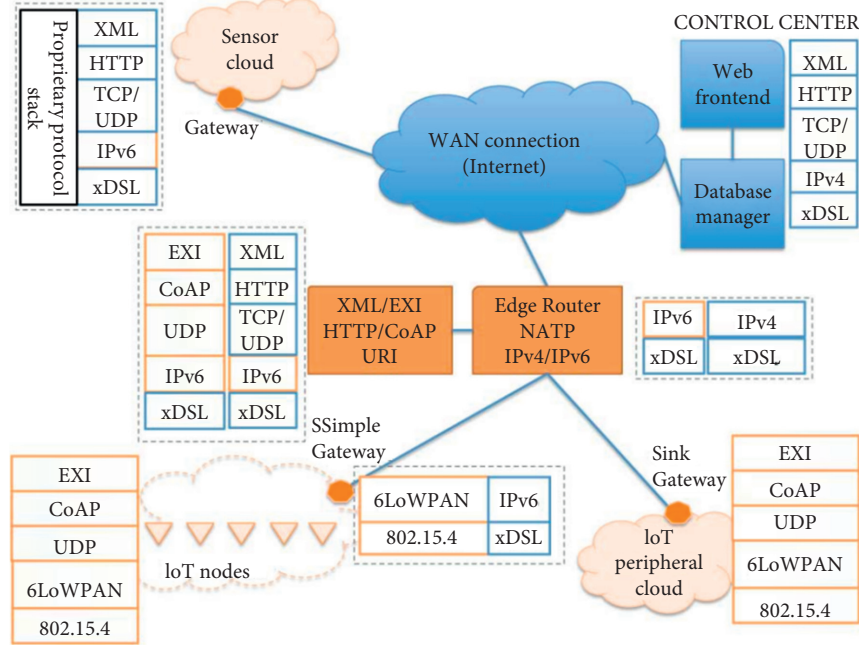
FIGURE 6: Conceptual image of a city IoT network.

$$n = \sqrt{i + o} + m. \qquad (6)$$

In the formula, n, i, and o, respectively, represent the number of monitoring points, the number of calculation units, and the number of display units.

3.3. Overall System Scheme and Basic Software and Hardware Design. From Figure 7, we can see that the building environment monitoring system mainly includes three modules: the collection and control module of environmental parameter information, the information management center module, and the host computer or user terminal module.

(1) The collection and control module of environmental information adopts *ZigBee* technology to form a wireless communication network. The *ZigBee* terminal node is connected to a sensor that collects environmental information and a controller that adjusts and controls the environment. The sensor periodically collects environmental data. The *ZigBee* terminal node sends the collected environmental data to the *ZigBee* coordinator node, which is responsible for collecting the environmental data collected by the terminal node, and then transmits it to the *ARM* home gateway. At the same time, the coordinator node is also responsible for forwarding the control instructions of the home gateway to the terminal node to realize the process of automatic control.

(2) The information management center module uses the *ARM* embedded server as the home gateway, stores the environmental parameters in the database, processes and analyzes the collected environmental information according to the preset environmental

parameters, and automatically sends the control instructions to the corresponding front-end controller. *ARM* embedded low-power and high-performance features make the system more secure, reliable, and energy efficient.

(3) User terminals mainly include communication equipment such as smart phones, tablet computers, and *PCs*. These devices can be directly connected to the gateway through Wi-Fi indoors, view environmental parameters, and send operating instructions, while outdoors they need to be remotely controlled through a *WEB* server.

In order to comprehensively monitor environmental information, the system detects and controls environmental parameters as follows: Temperature and humidity are the most sensitive factors affecting human comfort. The main components of household gas carbon monoxide and formaldehyde gas produced by furniture and decoration are common indoor pollutant gases, which have great harm to people's health. When the set threshold is exceeded, the system maintains air circulation by controlling the ventilation equipment and activates the alarm device (Figure 8). Fire and illegal intrusion are related to the safety of people's lives and properties. When a fire information or illegal intrusion event is detected, the system will immediately activate the fire or security alarm system.

As the core of the whole system, the home gateway and project center process the environmental data and judge the control mode after reading the environmental data. If it is in the manual control mode, it will enter the manual control strategy, and carry out environmental control according to the user's control instructions. If it is an automatic control mode, the judgment of the scene mode is carried out according to the data of the environment, and each scene
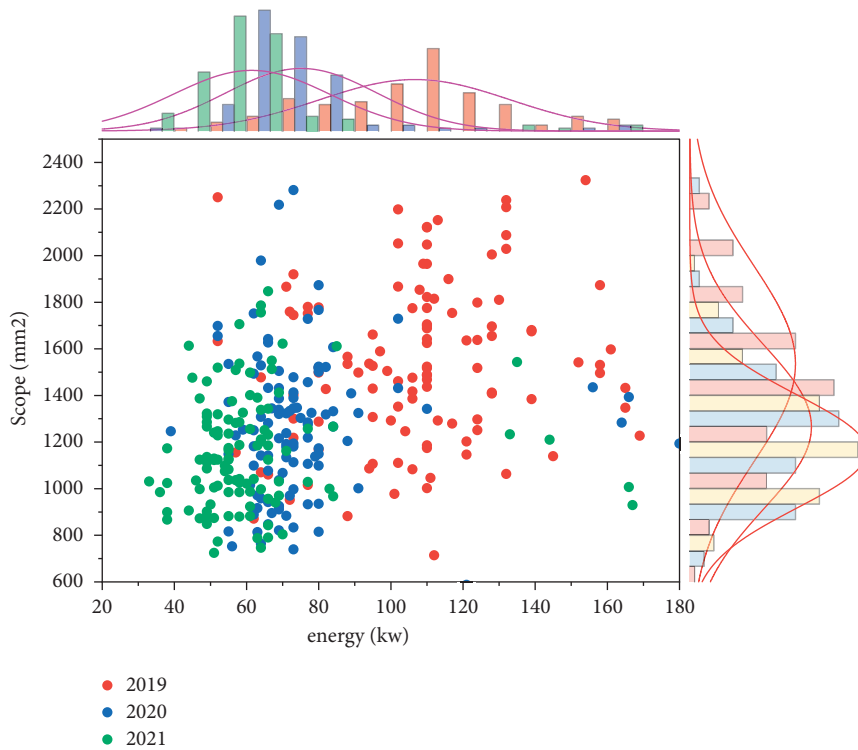
FIGURE 7: Harmful gas movement range and release energy.
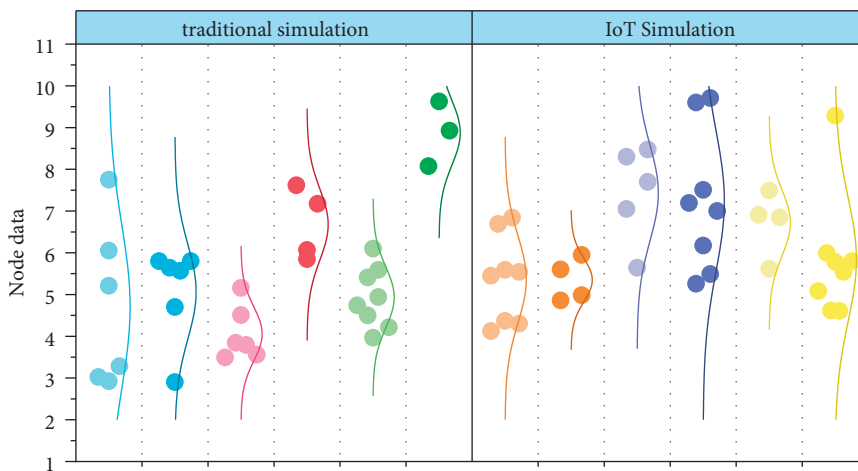
FIGURE 8: Various parameter values for the two algorithms.

mode will correspond to its corresponding control strategy and store and display the data.

According to the characteristics of the intelligent environmental monitoring system, the principle of modularity should be adopted in the design of the system, so that each module has its independence. In this way, when the system adds new modules or exits old modules, the system will remain stable. At the same time, since each sensor node needs to work for a long time, it is necessary to reduce the power consumption of the node as much as possible. In order to reliably realize the functions of the intelligent environmental monitoring system, it is necessary to design from two aspects of the system's hardware and system's software.

4. System Hardware Design

The system uses *ZigBee* technology to monitor the environmental parameters of the building site. First, use the terminal sensor node to collect environmental data, and then send it to the coordinator terminal node through *ZigBee* two-way wireless communication, upload it to the remote client, and display it in real time through the host computer. The user can monitor the site environment in real time through the host computer, and if an abnormality is found, the alarm can be timely and the intelligent equipment inside the building can be reversely controlled to deal with the dangerous situation. In order to facilitate the early detection

of abnormalities by the personnel in the building, relevant environmental parameters will be directly displayed through the *LCD* liquid crystal display installed on the coordinator node.

$$V_{ij} = X_{ij} + \text{rand}(-1,1)(X_{ij} - X_{kj}). \tag{7}$$

Among them, i and j represent the objective function value of the ith node, and k is a random number.

ZigBee technology is one of the wireless communication technologies. Compared with wired transmission, it has great advantages in construction sites with high security requirements. It has the advantages of low-power consumption, low cost, short distance, low complexity, and fast connection speed. In addition, *ZigBee* communication connection takes only 25 ms, while other wireless communication technologies such as *Bluetooth* require 5 s and Wi-Fi takes 11 s. In the face of frequent environmental data transmission, *ZigBee* is more in line with the project requirements.

$$P_i = \frac{f(X_i)}{\sum_{n=1}^{N_s} f(X_n)}, \tag{8}$$

where $f(x)$ is the software and environment fitness value.

In this design, a single-chip integrated *SoC* solution is mainly used, which is mature, small in size, and convenient and flexible to use. The *ZigBee* hardware platform selects the *CC2530* chip produced by *TI*. Although the sleep voltage of the chip is not the highest, its transmit power and receive sensitivity are the highest, reaching 4.5 dBm and −97 dBm, respectively.

$$Y = Hx + [e_1, e_2, e_3, \ldots, e_n]^T. \tag{9}$$

With the ambient noise and the sensor noise, the actual measured value of the system is x.

This system needs to realize real-time data uploading to the cloud, so the system chooses the *ESP-12F* Wi-Fi module developed by Anxinke Technology. The module is inexpensive and mature in technology, with an operating temperature range of −40 to 125°C. The temperature setting can ensure the stability of the environment in the construction site, with low standby power and fast wake-up speed, which can meet the requirements of real-time data transmission.

$$\hat{X} = (H^T W H)^{-1} H^T W Y, \tag{10}$$

where W is a positive definite diagonal weighting matrix.

Use *DHT11* sensor to collect warehouse temperature and humidity data. *DHT11* collects and combines temperature and humidity data, outputs it in the form of digital signal, and uses special analog signal acquisition, conversion technology, and temperature and humidity sensing technology to ensure that the sensor has good stability and reliability. The average working current of *DHT11* does not exceed 0.5 mA, the power consumption is low, and the working voltage range is 3~5.5 V, which matches the power supply equipped in this design.

5. System Software Design

Sensor-related operations include initializing the sensor, reading sensor values, and writing data to an array in a fixed format. During initialization, it is only necessary to open the clock of the corresponding port and initialize the *GPIO*. The initialization of the environmental quality detection sensor and the *PM2.5* sensor needs to initialize the corresponding *UART*. The former baud rate is 9600; the latter is 2400. From Figure 9, when the initialization is completed, the *UART* needs to be enabled, the Interrupt Enable is turned on, and the first data acquisition after power-on is performed. Finally, the obtained data is filled into the corresponding array according to the node type, and the reading of the environmental quality monitoring sensor refers to the fixed data format.

Wi-Fi mainly includes operations such as initializing the module, sending basic commands, and sending and receiving messages. This system uses *ESP8266* chip for network operation; the chip can set up *TCP/IP* server and monitor automatically through serial port commands. If a remote connection is established with the chip through *Socket* communication, the server will automatically receive this request and automatically assign a source *ID* for communication with the current connection. After the connection is established, if there is still no actual communication content after a certain period of time, the connection will be automatically disconnected to prevent the connection line from being occupied for a long time.

$$E[(x - \hat{x})^2] = E\left\{\sum_{i=1}^{n}\left[\left(\frac{w_i}{\sum_{i=1}^{n} w_i}\right)\right]^2 \varepsilon_i^2\right\}. \tag{11}$$

The above formula represents the sum of the squares of the estimation errors of x.

Select a residential building as an example in the design, import the model into the Unity engine, and add text boxes for displaying data. As Figure 10 shows, a text box for displaying data related to environmental quality is preset on the user form in advance, and the text box on the interface is bound to the corresponding control variable in the main program through the object binding process. When the application is running, a query command is sent to the designated environmental quality monitoring node server, and the text box is filled by parsing the returned data; that is, real-time data embedding is achieved.

UML provides a variety of model diagrams. In software development, these model diagrams can be used to describe the work in each stage of development, which is convenient for work implementation and personnel communication. In real-world applications, *UML* is mainly used in two phases: requirement analysis and system design. In the requirements analysis stage, the use case model diagram is mainly used to describe the system requirements more intuitively and facilitate communication with users. In the design phase of the system, the static model and dynamic model of the system are mainly used.

$$\varepsilon_{mi}^2 = E(y_{mi} - \overline{y_m})^2, \tag{12}$$
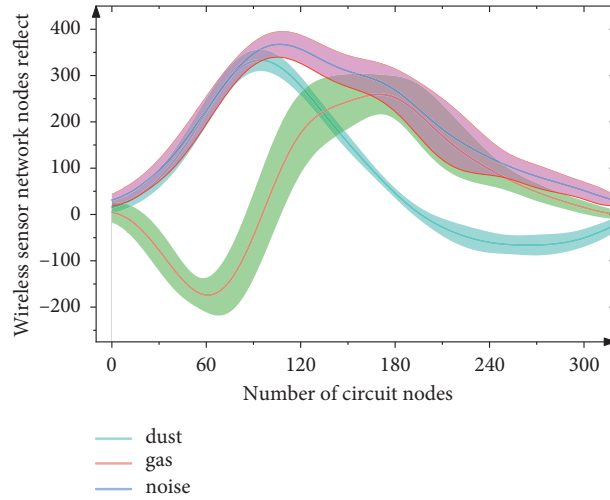
FIGURE 9: Wireless sensor circuit network node unit diagram.
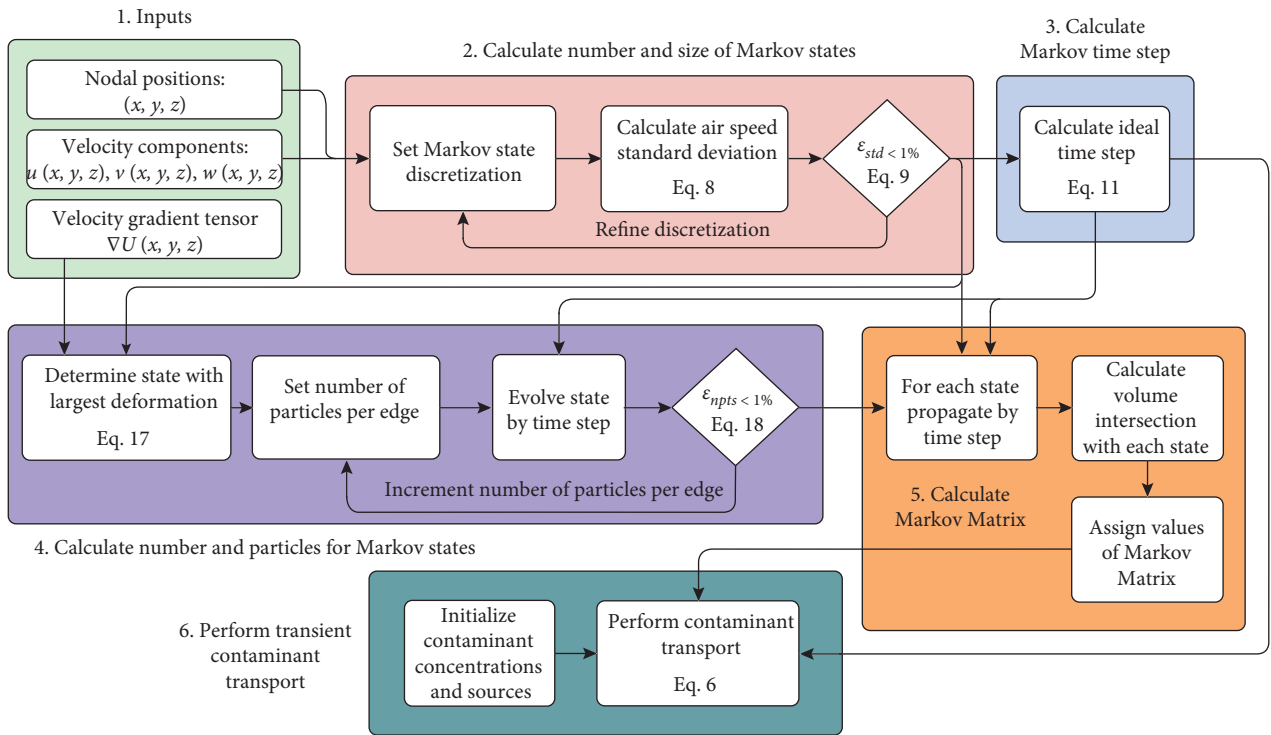
FIGURE 10: Method from input to accurate matrix of contaminant transport.

where ε is the measurement variance of the mth measurement from the ith sensor.

6. System Testing and Evaluation

In order to test the function of the system, temperature and humidity, harmful gases, and illegal intrusion are used as test samples, so that different environmental areas in the building can be simulated. The ambient temperature and humidity are detected by the *DHT11* temperature and humidity sensor with strong anti-interference ability. Hazardous gas sensors use *MQ-5* semiconductor combustible gas sensors with high sensitivity, and *HC-SR501* human body infrared sensors are used to monitor illegal intrusion events. The biggest advantage of smart site construction project construction is that the site staff can remove faults in time and reduce the occurrence of engineering project safety accidents.

Start each module of the system, the environmental data acquisition system will send the network access request to the gateway through the *ZigBee* wireless sensor network, and the gateway will send the command to assign *ID* after receiving the request instruction. If the light-emitting diode of the environmental data acquisition system flashes three times at this time, it means that the environmental data acquisition system is successfully connected to the network.

In this test, the on/off control of the relay is simulated by the brightness of the *LED* light.

Compared with the traditional environmental monitoring system, this system has the following advantages:

(1) Simple and scalable. There is no need to specially set up independent systems for different objects. In this system, multiple different environmental objects can be monitored, transmitted, and processed in the same network.

(2) The accuracy of the data is high. In the *ZigBee* network, the adopted intelligent optimization algorithm makes comprehensive decision-making, which avoids the error caused by monitoring a single node, and can greatly improve the accuracy of data monitoring.

(3) The node operation is flexible. Users can add environmental information collection nodes or reduce collection nodes according to their own needs, without requiring additional wiring, and the nodes are flexible in operation and can be placed freely.

7. Conclusion

In this paper, through the analysis of the research status of environmental monitoring system design and application scheme, a set of integrated real-time monitoring system based on Internet of Things technology has been successfully constructed, and it is truly stable in the environmental monitoring system, with physical measurement experiments for indoor and outdoor environments of buildings. Based on the scalability and wide application of the system, scholars in the related fields of construction engineering science in the future can expand as needed to meet the comprehensive real-time monitoring of various building types. The platform connects to the Internet through Wi-Fi wireless network for data communication, which can realize office or home networking, and real-time sharing of monitoring data from multiple monitoring points. The access and setting of the equipment can be completed through the computer or mobile *APP*, and the display software is used to realize the three-dimensional visualization of the indoor environmental quality parameter information in the building model. Internet of Things technology has been continuously accumulated and upgraded, and the industrial chain has gradually improved and matured. IoT technology will bring opportunities to upgrade infrastructure construction around the world.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] A. Esteves, M. J. Esteves, M. V. Mercado, G. Barea, and D. Gelardi, "Building shape that promotes sustainable architecture. Evaluation of the indicative factors and its relation with the construction costs," *Architecture Research*, vol. 8, no. 8, pp. 111–122, 2018.

[2] G. Marques and R. Pitarma, "An indoor monitoring system for ambient assisted living based on internet of Things architecture," *International Journal of Environmental Research and Public Health*, vol. 13, no. 11, p. 1152, 2016.

[3] M. Aksoezen, M. Daniel, U. Hassler, and N. Kohler, "Building age as an indicator for energy consumption," *Energy and Buildings*, vol. 87, pp. 74–86, 2015.

[4] G. Ala, G. Di Filippo, F. Viola et al., "Different scenarios of electric mobility: current situation and possible future developments of fuel cell vehicles in Italy," *Sustainability*, vol. 12, no. 2, p. 564, 2020.

[5] C. N. Antoniadis and G. Martinopoulos, "Optimization of a building integrated solar thermal system with seasonal storage using TRNSYS," *Renewable Energy*, vol. 137, no. 1, pp. 56–66, 2019.

[6] M. Ge, H. Bangui, and B. Buhnova, "Big data for internet of Things: a survey," *Future Generation Computer Systems*, vol. 87, pp. 601–614, 2018.

[7] M. Andronie, G. Lăzăroiu, M. Iatagan, C. Uţă, R. Ştefănescu, and M. Cocoşatu, "Artificial intelligence-based decision-making algorithms, internet of Things sensing networks, and Deep learning-assisted smart process management in cyber-physical production systems," *Electronics*, vol. 10, no. 20, p. 2497, 2021.

[8] G. Arslan, "Loneliness, college belongingness, subjective vitality, and psychological adjustment during coronavirus pandemic: development of the College Belongingness Questionnaire," *Journal of Positive School Psychology*, vol. 5, no. 1, pp. 17–31, 2021.

[9] C. K. Ng, C. H. Wu, K. L. Yung, W. H. Ip, and T. Cheung, "A semantic similarity analysis of Internet of Things," *Enterprise Information Systems*, vol. 12, no. 7, pp. 820–855, 2018.

[10] L. Lei, Y. Tan, K. Zheng, S. Liu, K. Zhang, and X. Shen, "Deep reinforcement learning for autonomous internet of Things: model, applications and challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1722–1760, 2020.

[11] M. Harrer, S. H. Adam, R. J. Fleischmann et al., "Effectiveness of an internet-and app-based intervention for college students with elevated stress: randomized controlled trial," *Journal of Medical Internet Research*, vol. 20, no. 4, Article ID e9293, 2018.

[12] J. J. P. C. Rodrigues, D. B. De Rezende Segundo, H. A. Junqueira et al., "Enabling technologies for the internet of health Things," *Ieee Access*, vol. 6, pp. 13129–13141, 2018.

[13] N. Limones-Rodríguez, J. Marzo-Artigas, M. F. Pita-Lopez, and M. P. Díaz-Cuevas, "The impact of climate change on air conditioning requirements in Andalusia at a detailed scale," *Theoretical and Applied Climatology*, vol. 134, pp. 1047–1063, 2018.

[14] N. Lolli and A. G. Hestnes, "The influence of different electricity-to-emissions conversion factors on the choice of insulation materials," *Energy and Buildings*, vol. 85, pp. 362–373, 2014.

[15] L. M. López-Ochoa, J. Las-Heras-Casas, L. M. López-González, and C. García-Lozano, "Environmental and energy impact of the EPBD in residential buildings in cold Mediterranean zones: the case of Spain," *Energy and Buildings*, vol. 150, pp. 567–582, 2017.

[16] G. De Luca, I. Ballarini, A. Lorenzati, and V. Corrado, "Renovation of a social house into a NZEB: use of renewable energy sources and economic implications," *Renewable Energy*, vol. 159, pp. 356–370, 2020.

[17] S. Grimmelikhuijsen, S. Jilke, A. L. Olsen, and L. Tummers, "Behavioral public administration: combining insights from public administration and psychology," *Public Administration Review*, vol. 77, no. 1, pp. 45–56, 2017.

[18] D. Zilio, "On the autonomy of psychology from neuroscience: a case study of skinner's radical behaviorism and behavior analysis," *Review of General Psychology*, vol. 20, no. 2, pp. 155–170, 2016.

[19] W. Wood and D. Rünger, "Psychology of habit," *Annual Review of Psychology*, vol. 67, no. 1, pp. 289–314, 2016.

[20] P. Sheeran, J. A. Bosch, G. Crombez et al., "Implicit processes in health psychology: diversity and promise," *Health Psychology*, vol. 35, no. 8, pp. 761–766, 2016.

Research Article

Online Shopping Brand Sales Based on IoT Big Data Processing

Menglin Zhang [iD][1] **and Xiaoyu Ma**[2]

[1]*Department of Fundamentals, Jiangsu Vocational College of Finance and Economics, Huaian 223003, Jiangsu, China*
[2]*School of International Education, Jiangsu Maritime Institute, Nanjing 210000, Jiangsu, China*

Correspondence should be addressed to Menglin Zhang; 19900514@jscj.edu.cn

As a pair of closely related technical fields, "Internet of Things" and "big data" are gradually applied to the analysis of online shopping, which is very beneficial to promote the healthy and orderly development of the online shopping industry. This article uses the big data processing technology of the Internet of Things to analyze the influencing factors of online shopping brand sales based on the data collected by online questionnaires. It draws certain rules from it, and provides relevant marketing strategies for online shopping brand sales to assist its market positioning. According to the survey, the five aspects of online evaluation, brand awareness, online shopping community, online public opinion influence of related brands, and other brand information acquisition channels all have an impact on online brand sales. Especially women aged 25–35 will pay attention to online information, and 195 people will not choose to buy from brands with negative news. And during the purchase process, 186 people will check the online reviews of related brand products. Among them, there are 183 men aged 25–35 who will pay attention to online reviews.

1. Introduction

According to CNNIC statistics, as of 2020, the number of search engine users reaches 790 million, and the utilization rate among netizens reaches 82.5%, as shown in Table 1. The popularity of the Internet has not only enabled people to share and obtain information, but also shaped a brand-new business model. Research shows that companies that choose e-commerce platforms for promotion account for the largest proportion, reaching 25.2%. Massive Internet search data records people's hobbies and needs, and provides a data foundation for social and economic development. Online shopping has become the mainstream mode of consumption in today's society. The selection of specific online marketing methods for small and medium-sized enterprises in 2020 is shown in Table 2.

(1) Theoretical significance: through the search of relevant information, most of the research on online shopping uses statistical analysis methods and almost no one cares about the big data processing technology of the Internet of Things. This article discusses the influencing factors of online shopping brand sales through data preparation, Internet of Things big data processing technology, result analysis, discussion and induction. It provides a reference for the research of related online sales and e-commerce industry in the future.

(2) Practical significance: this article collects relevant data through questionnaires, and uses the Internet of Things big data processing technology to analyze, so as to find some rules in online brand shopping. This provides marketing reference for companies engaged in online item sales and operations, and at the same time provides some theoretical foundations for the management of online platforms.

With the rapid development of the Internet of Things technology, many application fields related to the Internet of Things have emerged [1], such as GPS-based Internet of Vehicles systems, Internet of Things logistics systems based on RFID technology, as well as applications in the medical industry, and so on. However, relevant research and analysis on online shopping using the big data processing of the

TABLE 1: Use of search engines by Chinese netizens.

Years	Data	
	Chinese search engine user scale/ten thousand people	Utilization rate among netizens (%)
2017	63956	82.8
2018	68132	82.2
2019	72106	81.8
2020	79461	82.5

TABLE 2: Selection of specific online marketing methods for SMEs in 2020.

Corporate internet marketing	The proportion (%)
E-commerce platform promotion	25.2
E-mail marketing	23.5
Search keyword ads	15.3
Internet brand advertising	13.2
Online video advertising	9.3
Network soft article	8.7
Search engine optimization	2.5
Network alliance, joint marketing	2.3
Other online marketing activities	1.1

Internet of Things through the search of relevant information remain scarce. Therefore, this article will analyze the influencing factors of online shopping brand sales through the Internet of Things big data processing technology. It not only promotes the application research of the big data processing technology of the Internet of Things, but also provides certain market analysis value for the e-commerce field.

2. Related Work

With the advent of the Internet age, online shopping has become a mainstream shopping model. With the prosperity of the consumer market and the intensification of commodity competition, the producer-led marketing concept is clearly no longer in line with market rules, and more and more businesses are beginning to pay attention to consumers. Therefore, it is an opportunity to use the big data processing technology of the Internet of Things to analyze the influencing factors of online merchandise sales. Internet of Things (IoT) related applications have become an important area for engineers and researchers. Cai H provides a functional framework for identifying the acquisition, management, processing and mining of big data in the Internet of Things, and defines and describes several related technical modules based on their key characteristics and capabilities. He also analyzes the current research on the application of the Internet of Things and identifies the challenges and opportunities related to the research on the Internet of Things big data [2]. Big data refers to a collection of massive amounts of data that cannot be processed by traditional data processing tools and technologies. A large amount of climate data is collected from IoT weather sensor devices and NCEP. Manogaran G proposes a big data processing framework to integrate climate and health data to find the correlation between climate parameters and the

incidence of dengue fever. The MapReduce framework based on In-Mapper is used to calculate the monthly average of the climate parameters for each latitude and longitude. Experimental results prove that, compared with the existing MapReduce algorithm, the In-Mapper based combiner algorithm has lower response time effectiveness [3]. The Internet of Things generates a lot of big data because it contains billions of devices connected to each other via the Internet. In order to process massive real-time geospatial big data, Limkar and Jha explores a scalable and efficient indexing method. A large number of experimental results show that the generated R-tree and its variants retain similar characteristics to the sequentially generated R-tree and its variants. This has excellent scalability, and reduces a lot of time to build indexes, index updates, and perform spatial range queries [4]. In recent years, e-commerce has been expanding at an alarming rate, and relevant Internet pages called e-services capes now occupy an important position in the business world. Therefore, Wu et al. conducts research on describing the nature of electronic service scenarios, and investigates the relationship between website credibility, website attitude, brand attitude, electronic word-of-mouth intention and purchase intention. A total of 290 responses are collected from Taiwanese consumers through online questionnaires. His research results show that the dimensions of the e-service landscape (aesthetic appeal, customization, usability, and financial security) have a significant impact on consumers' attitudes and trust in websites [5]. Ladhari et al. has also done related research. The goal of his research is to subdivide Gen female online shoppers based on their psychological, demographic and behavioral characteristics. The data comes from consumers of a female fashion retailer, and a total of 538 women shopping on this fashion retailer's website completed the survey. The results reveal four methods of online shopping: trend shopping, fun shopping, price shopping and brand shopping. It also identifies six shopping profiles, each with different goals: price shoppers, discovery shoppers, emotional shoppers, strategic shoppers, fashionistas and shopping fans [6]. In addition, Abbes I and others have studied to determine the stakes of the brand collaborative redistribution platform, and understand the influence of its internal and external characteristics on behavioral intentions. He also conducts a quantitative study on 214 people who had purchased second-hand goods online. The results show that the loyalty intention of the collaborative redistribution platform has an influence on the loyalty intention of the brand. The influence of platform service experience satisfaction on brand loyalty intention is mediated by platform loyalty intention [7]. Swapana and Padmavathy also conducts related research to

study the relationship between the three dimensions of online second-hand shopping, customer satisfaction and repurchase intention. The responses received from 608 Indian online second-hand shoppers indicate that the dimensions of online second-hand shopping, such as price, website quality, nostalgia, and brand image, affect customer satisfaction. In addition, customer satisfaction has a positive impact on repurchase willingness. The research results show that online second-hand marketers enter emerging markets through effective online shopping strategies and provide intuitive guidance [8]. It can be seen that with the rise of online shopping, in the current era of large scale use of IoT big data to achieve efficient data processing, a decided lack of literature on the sales of online shopping brands remains. Therefore, in order to better analyze the influencing factors of online shopping brand sales, this article will use IoT big data processing technology. This has also promoted the further development of the Internet shopping era.

3. Internet of Things Big Data Processing Technology

Apriori algorithm and clustering algorithm are currently the two most widely used data mining algorithms [9, 10]. Its main research fields have a wide range of applications from pattern recognition to data analysis and image processing. Especially for the processing of big data in the Internet of Things, its characteristics are excellent, thus preferred by academic research and widely used in the industry [11, 12].

3.1. Clustering Algorithm. Two Step is a kind of clustering algorithm, which is used to solve the cluster analysis problem of massive data and complex category structure [13]. The difference between this algorithm and the traditional clustering algorithm is that the two types of clustering methods have small memory and fast calculation speed. Two-step clustering is to complete the clustering work through the two steps of preclustering and formal clustering [14, 15]. In the preclustering stage, data points in the dataset need to be inserted into the clustering feature tree (CF tree) one by one to realize the growth of CF tree. When the volume of the CF tree exceeds the set size, the CF tree needs to be thinned. On the contrary, the outliers of the thin CF tree are inserted into the CF tree. After reading all the data points, the potential outliers that cannot be inserted into the CF tree are the true outliers. Finally, the clustering features of the corresponding subclusters of the final CF leaf element are output to the next stage of the algorithm. As shown in from formulas (1) to (9) [16, 17].

$$\vec{\tilde{\Lambda}}_j = \left(\tilde{\Lambda}_{j1}, \ldots, \tilde{\Lambda}_{jD_1} \right)^T = \left(\sum_{n=1}^{N_j} \tilde{x}_{jn1}, \ldots, \sum_{n=1}^{N_j} \tilde{x}_{jnD_1} \right)^T, \quad (1)$$

$$\vec{\tilde{\Sigma}}_j = \left(\tilde{\Sigma}_{j1}, \ldots, \tilde{\Sigma}_{jD_1} \right)^T = \left(\sum_{n=1}^{N_j} \tilde{x}_{jn1}^2, \ldots, \sum_{n=1}^{N_j} \tilde{x}_{jnD_1}^2 \right)^T, \quad (2)$$

where $\vec{\tilde{\Sigma}}_j$ represents the linear summation of the attribute values of the data points in cluster $C_{-}j$.

$$\vec{N}_j = \left(\left(\vec{N}_{j1} \right)^T, \ldots, \left(\vec{N}_{jD_2} \right)^T \right)^T, \quad (3)$$

where $\vec{N}_j$ represents the sum of the squares of the attribute values of the data points in cluster $C_{-}j$.

$$\vec{CF}_{\langle j, j' \rangle} = \langle N_j + N_{j'}, \vec{\tilde{\Lambda}}_j + \vec{\tilde{\Lambda}}_{j'}, \vec{\tilde{\Sigma}}_j + \vec{\tilde{\Sigma}}_{j'}, \vec{N}_j + \vec{N}_{j'} \rangle, \quad (4)$$

where $\vec{CF}_{\langle j, j' \rangle}$ represents clustering characteristics.

$$\zeta_j = -N_j \left(\frac{1}{2} \sum_{s=1}^{D_1} In \left(\hat{\sigma}_{js}^2 + \hat{\sigma}_s^2 \right) + \sum_{t=1}^{D_2} \hat{E}_{jt} \right), \quad (5)$$

where ζ_j represents the definition parameter of cluster $C_{-}j$.

$$\hat{\sigma}_{js}^2 = \frac{1}{N_j} \sum_{n=1}^{N_j} \left(\tilde{x}_{jns} - \tilde{x}_{js} \right)^2, \quad (6)$$

$$\hat{\sigma}_{js}^2 = \frac{\tilde{\Sigma}_{js}}{N_j} - \left(\frac{\tilde{\Lambda}_{js}}{N_j} \right)^2, \quad (7)$$

$$\hat{\Sigma}_{jt} = -\sum_{k=1}^{\epsilon_t} \left(\frac{\ddot{N}_{jtk}}{N_j} In \frac{\ddot{N}_{jtk}}{N_j} \right)$$
$$= -\frac{\vec{\ddot{N}}_{jt}^T}{N_j} \cdot In \frac{\vec{\ddot{N}}_{jt}}{N_j} - \left(1 - \frac{\vec{1}^T \vec{\ddot{N}}_{jt}}{N_j} \right) \cdot In \left(1 - \frac{\vec{1}^T \vec{\ddot{N}}_{jt}}{N_j} \right), \quad (8)$$

$$d\left(C_j, C_{j'} \right) = \zeta_j + \zeta_{j'} - \zeta_{\langle j, j' \rangle}, \quad (9)$$

where $d\left(C_j, C_{j'} \right)$ represents the log-likelihood distance of cluster $C_{-}j$ and cluster $C_{-}\{j'\}$.

The input of the clustering stage is the subclusters of the leaf element items of the final CF tree output by the preclustering stage. Therefore the work of this stage is to perform two-degree clustering on the subclusters according to the input data, and finally achieve the clustering results of the expected number of clusters. As shown in from formulas (10) to (17) [18, 19].

$$BIC(M_i) = -2Inf_i\left(x; \hat{\theta}_i \right) + k_i Inn, \quad (10)$$

$$BIC(C_J) = -2L_{CJ} + KInN, \quad (11)$$

$$L_{CJ} = \sum_{j=1}^{J} L_{Cj} = \sum_{j=1}^{J} \zeta_j, \quad (12)$$

$$K = J \left(2D_1 + \sum_{t=1}^{D_2} \left(\epsilon_t - 1 \right) \right), \quad (13)$$

$$\Delta_{BIC}(J) = BIC(C_J) - BIC(C_{J+1}), \quad J = 1, \ldots, J_0, \quad (14)$$

where $\Delta_{BIC}(J)$ represents the difference between the BIC values of adjacent clusters.

$$r_1(J) = \frac{\Delta_{BIC}(J)}{\Delta_{BIC}(1)}, \tag{15}$$

where $r_1(J)$ represents the rate of change of BIC value;

$$r_2(J) = \frac{d_{\min}(C_J)}{d_{\min}(C_{J+1})}, \quad J = J_I, \dots, 2, \tag{16}$$

where $r_2(J)$ represents the ratio of the closest cluster distance of clusters C_J and $C\_\{J+1\}$.

$$J^* = \left\{ \begin{array}{ll} J_1, \text{if} & \frac{J1}{J} > 1.15 \\ \\ \max\{J_1, J_2\}, & \text{otherwise} \end{array} \right\}, \tag{17}$$

where J^* represents the best number of clusters selected from J_1 and J_2.

3.2. Apriori Algorithm.

Apriori algorithm is a classic data mining algorithm for mining frequent itemsets and association rules [20]. The algorithm can not only discover frequent itemsets, but also mine association rules between items. It uses support and confidence to quantify frequent itemsets and association rules. The algorithm uses the prior properties of frequent itemsets to compress the search space. As shown in the following formula [21, 22].

$$\text{support}(A \Rightarrow B) = P(A \cup B), \tag{18}$$

$$\text{confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{support}(A \cup B)}{\text{support}(A)}, \tag{19}$$

$$\text{confidence}(A \Rightarrow B) = \frac{\text{support\_count}(A \cup B)}{\text{support\_count}(A)}, \tag{20}$$

where

support$(A \Rightarrow B)$: it indicates the degree of support of the association rule

confidence$(A \Rightarrow B)$: it represents the confidence level of the association rule

4. Application of Big Data Processing Technology in the Internet of Things

The actual practice of big data in the Internet of Things shows that reasonable systems and methods play a decisive role in the process [23]. The association rule in the Apriori algorithm refers to the common relationship between items, which is usually used to analyze sales. The Two Step algorithm supports numerical and subtype data. For large amounts of data, the calculation speed is faster and the execution efficiency is high. It provides great convenience in the application [24, 25], so this article will use the Internet of Things big data processing technology to study the influencing factors of online shopping brand sales.

4.1. Data Preparation.

The survey method of this article is online questionnaire survey. It conducted 200 surveys on six groups of junior high school students, high school students, college students, young people aged 25–35, middle-aged people aged 35–50, and over 50 years old, and a total of nearly 200 questionnaire results for each group were recovered. At the same time, some face-to-face questionnaire surveys were conducted offline, and a total of 200 valid questionnaires were collected.

This article uses the Apriori algorithm and the two step algorithm to analyze the impact of online evaluation, brand awareness, online shopping community, related brand network public opinion influence, and other understanding channels on online brand sales, and conducts multiple fields. The results of this questionnaire survey are saved in an excel table, using its filtering function, selecting the T1 field (whether it will view and make relevant product reviews when shopping); selecting the T2 field (whether it will choose to buy products with high brand awareness); selecting the T3 field (whether it will buy through the online shopping community); selecting the T4 field (whether it will pay attention to and care about some negative news about the brand on the Internet) and selecting the T5 field (will it learn about the brand through other channels, such as the third website, offline word-of-mouth, etc.), and then filter out all records with a field value of "Yes". The composition in Figure 1 sorts out five factors that affect online brand sales, including online evaluation and brand awareness, These five aspects are based on the sources that affect online brand sales, and it can be divided into three categories: one category comes from the consumer's online activity itself, that is, online evaluation; one category comes from online brands, including brand awareness of shopping websites, online shopping communities and related network information. Although these three elements will be controlled by the brand, the online activities of customers also play an important role; the last category comes from other channels, including third-party websites and offline word-of-mouth communication. The source of these two elements is not obvious, but they actually exist in the Internet. At the same time, these two elements are also mixed with a large number of online customer activities.

4.2. Data Analysis.

The Internet continues to penetrate into the daily lives of netizens, and the number of online shopping users is increasing year by year. According to CNNIC statistics, as of 2020, the number of online shopping users in China has reached 989 million, and the Internet penetration rate has reached 70.4%. It is shown in Table 3. The development of commercial applications on the Internet is due to the rapid growth of the number of online shopping companies and that of Internet users who have begun to get
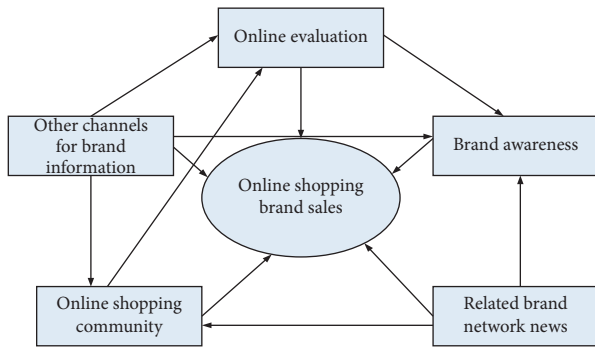
FIGURE 1: Based on the five aspects of the impact of online shopping brand sales composition.

TABLE 3: Statistical data table of Chinese internet development status.

| Years | Data | |
	Number of online shopping users	Internet penetration rate (%)
2017	7.72	55.8
2018	8.28	59.6
2019	9.03	64.5
2020	9.89	70.4

TABLE 4: Customer online evaluation questionnaire of major clothing brands in November.

| Clothing brand | Evaluation attributes | | |
	Praise	Average	Negative ratings
PRADA	425	29	3
ONLY	46	16	0
ZARA	52	18	16
HM	103	12	10

used to this new way of online shopping. Therefore, this article will conduct a quantitative analysis of online shopping brand sales from the following five aspects.

4.2.1. Online Evaluation. Online evaluation is a kind of online information, which mainly refers to the content itself spread by customers online. Every online behavior itself produces a large amount of information content, which will affect the popularity or reputation of online brands. Online information can often be analyzed from two aspects: information direction and information characteristics. Information direction refers to whether the information content itself has a positive or negative view of the online brand. Obviously, the direction of positive information can enhance the reputation of online brands, and vice versa, the direction of negative information will greatly damage the reputation of online brands. Information characteristics usually refer to the characteristics of the information, such as its manifestation, such as the length and type of the information, etc.; it also includes the inherent characteristics of the information, such as the credibility and legibility of the information. On the one hand, information characteristics will affect the information itself, and at the same time affect online brand sales through information.

As shown in Table 4, in November 2020, major apparel brands had a total of 730 online reviews by customers. Through analysis, PRADA pays more attention to online reviews of customers.

In the network context, the role played by the communicator and the receiver becomes difficult to distinguish, and the customer is both the communicator and the receiver. Generally speaking, the identity characteristics of the communicator and the receiver will obviously affect the

credibility of the information they provide, thereby affecting the sales of online brands. The first is the professionalism of the two, which specifically refers to whether the communicator or receiver is an expert in this field. In the network environment, because the identities of the two are difficult to identify, false information is inevitable. But the professionalism can be judged from the content of the information it provides, such as customer reviews. It provides information in the form of reasoning, arranging knowledge, and discussing opinions, which obviously can be convincing. The second is the authority of the two. Compared with professionalism, authority is more difficult to identify. In traditional media, the authority is easy to judge, such as the authenticity and credibility of mainstream media reports. However, in the network environment, it is difficult to determine the authority due to the difficulty of identification. However, as some network applications begin real-name systems, the authority of the communicator or receiver becomes easier to judge. For example, the real-name authentication of Weibo, celebrity Weibo or expert Weibo is more authentic and credible due to its authority. Obviously, if online communicators and receivers are professional or authoritative, the information they provide is more likely to affect online brand sales. Therefore, the survey results based on online evaluation are shown in.

The results of the male and female role survey also showed that, when most people shop, they check reviews and make comments on related product, as shown in Figure 2, the highest enthusiasm for the activity was women aged 25–35, with 186 people. Followed by female college students and 25–35-year-old males, 183 people, and 172 male college students will also have this activity. This shows that most people agree that online reviews have a great influence on their brand choice.

4.2.2. Brand Awareness. This article believes that brand awareness also occupies an important position, and its influence on online brand sales is second only to online customer reviews. Compared with offline shopping, online shopping has its own uniqueness, so many people pay more attention to the brand when buying things. The inability of consumers to directly touch and feel the product increases the difficulty of product quality judgment. It can be seen from this that in the highly competitive e-commerce market, brand awareness is the standard by which buyers' measure stores. The higher the popularity of the store, the natural flow of people will increase, and the sales will also increase. The result of the influence of brand awareness on brand sales is shown in Figure 3.
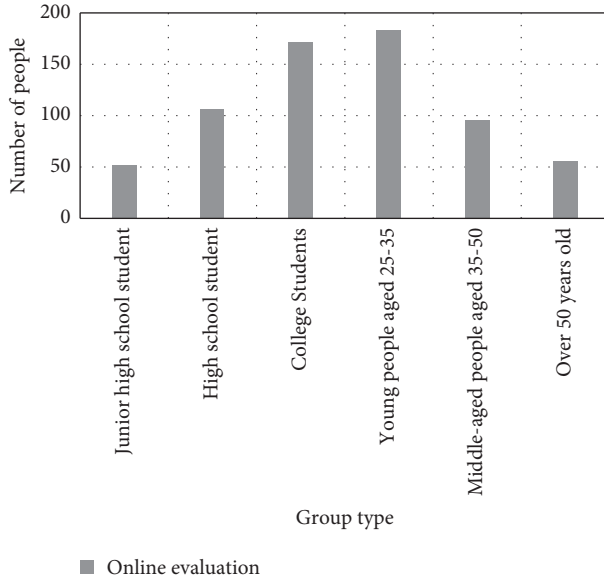
(a)
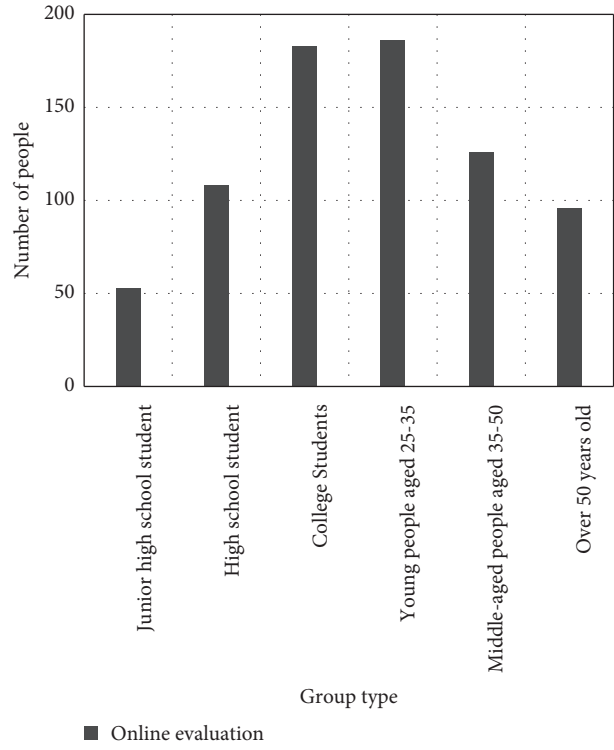
(b)

FIGURE 2: Online evaluation survey results. (a) Male survey results on online evaluation. (b) Women's survey results on online reviews.
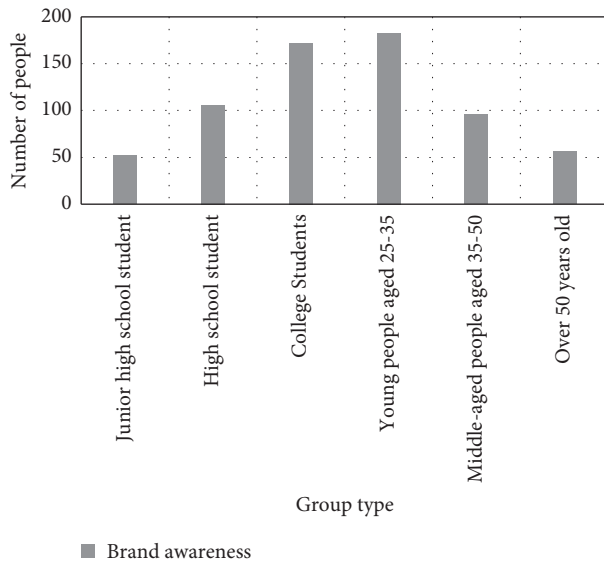
(a)
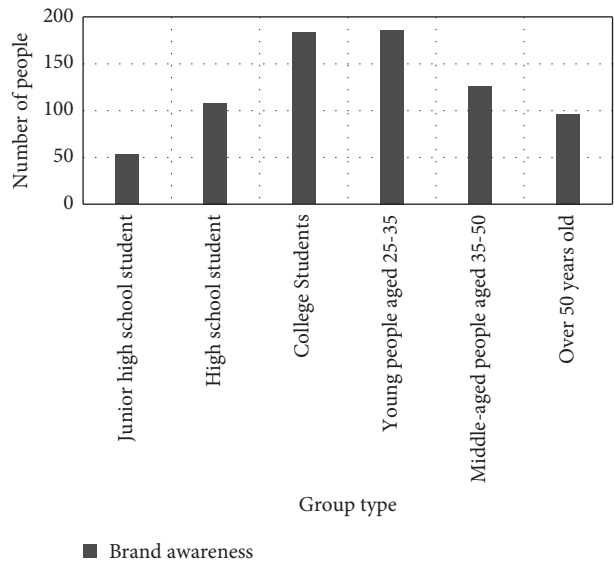
(b)

FIGURE 3: Survey results of brand awareness. (a) Men's survey results on brand awareness. (b) Women's survey results on brand awareness.

As mentioned above, brand awareness accounts for a large proportion of online brand sales. 132 male college students and 161 25–35-year-old males will choose well-known brands when shopping, and agree that the brand's products have certain quality assurance. As shown in Figure 3, female college students and women over the age of 25 trust the quality of well-known brands, and their numbers are above 110. This shows that brand awareness has a great influence on sales.

4.2.3. Online Shopping Community. Online shopping community generally refers to a special discussion space

built by merchants in order to activate the atmosphere of online shopping and enhance the initiative and enthusiasm of customers in online activities. Customers share shopping in the community, post orders and post product or service reviews. In the purchase process, users click on the product information of interest according to their own needs, and these evaluations can increase the attention of other customers to the product, thereby prompting customers to click and browse the product information. The results of the face-to-face questionnaire survey in this article show that 30.27%, 22.51%, and 17.25% of users agree that they will click on product information with higher attention, as shown in Figure 4. Not only that, the community homepage has a search function, which can help customers find their ideal products quickly and easily. It can be seen that the huge online activities in the online shopping community will undoubtedly involve all aspects of online brand sales. Therefore, this kind of combined customer online activities dominate online brand sales.

It can be said that the online shopping community is the essence of customer online activities. A variety of customer online activities gather in the online shopping community, such as the communication between the brand and the members, and that between community members. Relying on the contributions of relevant parties, these activities themselves have been continuously optimized. The information provided by community interaction is more comprehensive and specific, which can be said to be the essence of customer online activities. At the same time, the online shopping community itself has a judging function. For example, posts that are popular with customers tend to have a higher click-on rate, more information value, and more active participation among netizens. The refined customer online activities are more credible. Thus both the brand and the customers themselves are making contributions to customer online activities. For example, sellers will show better-selling products through online communities, while buyers will help latecomers through their own shopping experience. Of course, the information will undoubtedly improve the authority and professionalism of customers' online activities. This kind of information content through the discussion can summarize the whole picture of the problem better, thus making the customer's online activities more authentic and credible. Such authentic online activities will undoubtedly affect the sales of online brands. The results of the online shopping community's influence on brand sales are shown in Figure 5.

As shown in Figure 5, many young people are very active in online communities.

Regardless of gender, groups from junior high school students to 25–35 young people believe in product reviews in online communities. The number of them is more than 100, and the highest number is 25–35 year-old women, reaching 153. This shows that the marketing of online shopping communities has a positive role in promoting online brand sales.

4.2.4. Network Message. With the rise of the Internet, there is a huge amount of information on the Internet. In order to get more display opportunities for information, or to allow online brands themselves to gain more public opinion support, brands will inevitably hire network naval forces to spread information, thus creating a false scenario. Network naval forces belong to a relatively special group in the Internet. Generally speaking, the network navy is employed by the brand side, and its behavior is mostly controlled by the online brand side. In addition to the normal way of brand communication, the current online brand communication also has to rely on a large number of network navigators.

First of all, network naval forces generally disseminate information on a large scale, which greatly enriches online information. Although the quality of information varies from good to bad, it often has an absolute advantage in quantity. In terms of the nature of online information provided by the network navy, there are currently two main types: positive online information and negative online information. Positive online information is mostly to promote the enthusiasm of the brand itself or the product, thereby enhancing the popularity and reputation of the online brand. Negative online information is generally done by competitors, who use the network navy to publish information that is unfavorable to the brand itself or the product, and at the same time use malicious posts to attract the public's attention and combat competitors. It is not difficult to see from the above that a large amount of information created by the network navy makes people difficult to distinguish between true and false, and this online information will also affect people's evaluation of the brand.

Secondly, the network navy's all-round online promotion behavior. The promotion behavior of the network navy is often planned and implemented by the public relations company for three-dimensional promotion. There are five common situations in the online promotion of network navy: First, community promotion. It provides dissemination services in communities and forums. Relying on the requirements of the brand or related parties to formulate forum or community communication strategies and specific implementation rules, while targeting target consumers, implement forum precision marketing, and follow up services in accordance with the marketing plan. Second, hot speculation. It uses the focus of attention of the public, media, and netizens to follow up, and creates high-relevant communication topics, and controls and guides them in this process. Third, event marketing. It manufactures hot events related to online brands, carries out all-round packaging and hype, attracts the attention of relevant parties, and at the same time attracts follow-up reports from the media to a certain extent, extends the influence of communication, in order to achieve the goal of online brand marketing. For example, the event marketing of Ganlu has become the target of media hype to a certain extent. Fourth, Weibo and blog marketing. It uses Weibo, blogs, etc. To spread products or brands. There are mainly the following three forms: well-known blog marketing, self-built blog marketing and group-built blog marketing. These three forms are usually implemented in combination. Fifth, deleting negative news. It contacts search engines, uses artificial means or a combination of other illegal means to instantly obtain negative
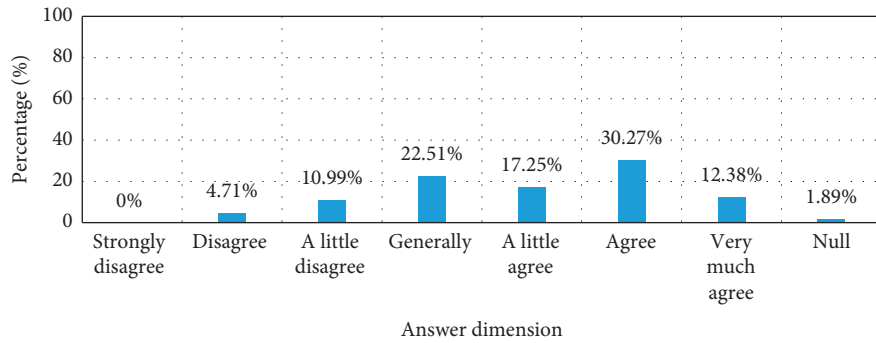
FIGURE 4: Product information or evaluations with a high degree of attention.
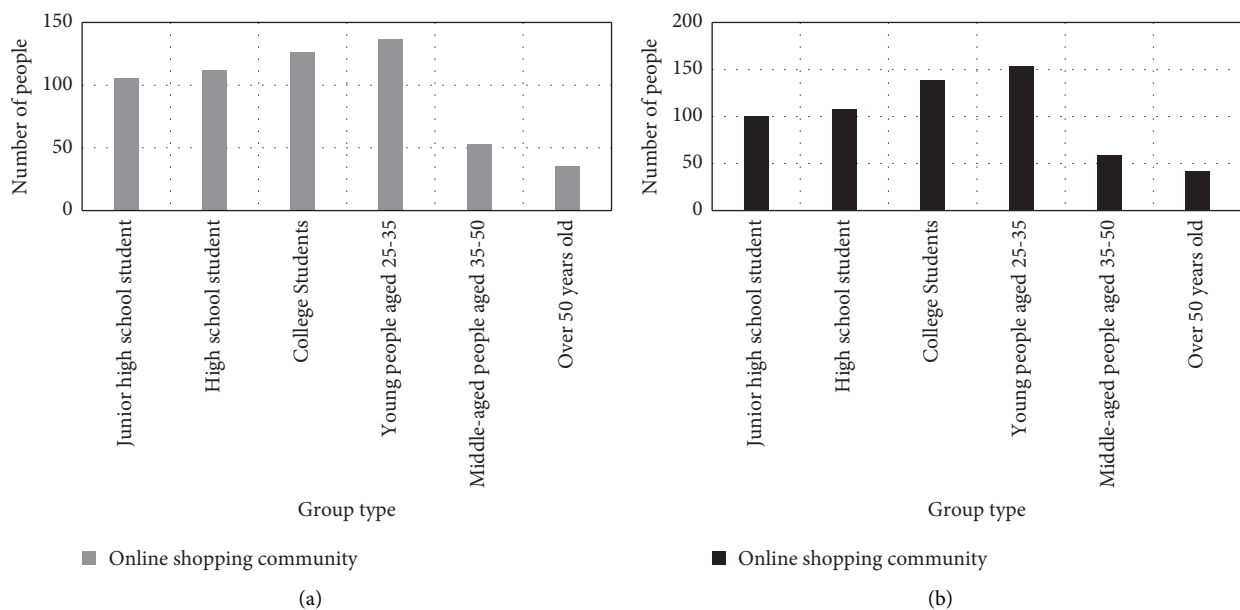
FIGURE 5: Survey results of online shopping communities (a) Male survey results on online shopping communities (b) Women's survey results of online shopping communities.

information related to online brands in the Internet space, and guide or clear it, especially suitable for active defense or prevention of brands negative information. From the above, the online promotion of the network navy has a guiding effect on the network information of related brands, thereby affecting the sales of online brands. Therefore, this article conducted a survey based on the elements of network information, and the survey results are shown in Figure 6.

In the Internet age, people can learn all kinds of information quickly without using traditional media, which brings great convenience to people. Therefore, the brand side will also hire the network navy to create the network focus, attract people's attention, so as to enhance the awareness and reputation of its own brand. As shown in Figure 6, women aged 25–35 will pay attention to online information, but for brands with negative news, they will not choose to buy, and the number of women reaches 195. Similarly, 182 men aged 25–35 also agreed with this view, and 186 female college students also expressed support. This shows that the

orientation of positive online news has a certain boosting effect on brand sales.

4.2.5. Other Channels.
Analysis of elements based on other channels, including third-party platforms, offline word-of-mouth communication.

(1) Third-Party Platform. Third-party platforms refer to customer online activities that appear on other Internet platforms in addition to the brand's own online marketing channels. Third-party websites also gather a large number of customer online activities. Since customer online activities appear on third-party platforms, its authenticity is stronger, and it is easier to get customers' affirmation.

First, the online activities of customers on third-party platforms are displayed. The ways for customers to display their activities online on third-party platforms are mainly self-media such as communities, forums, blogs, Weibo, and

(a)

(b)
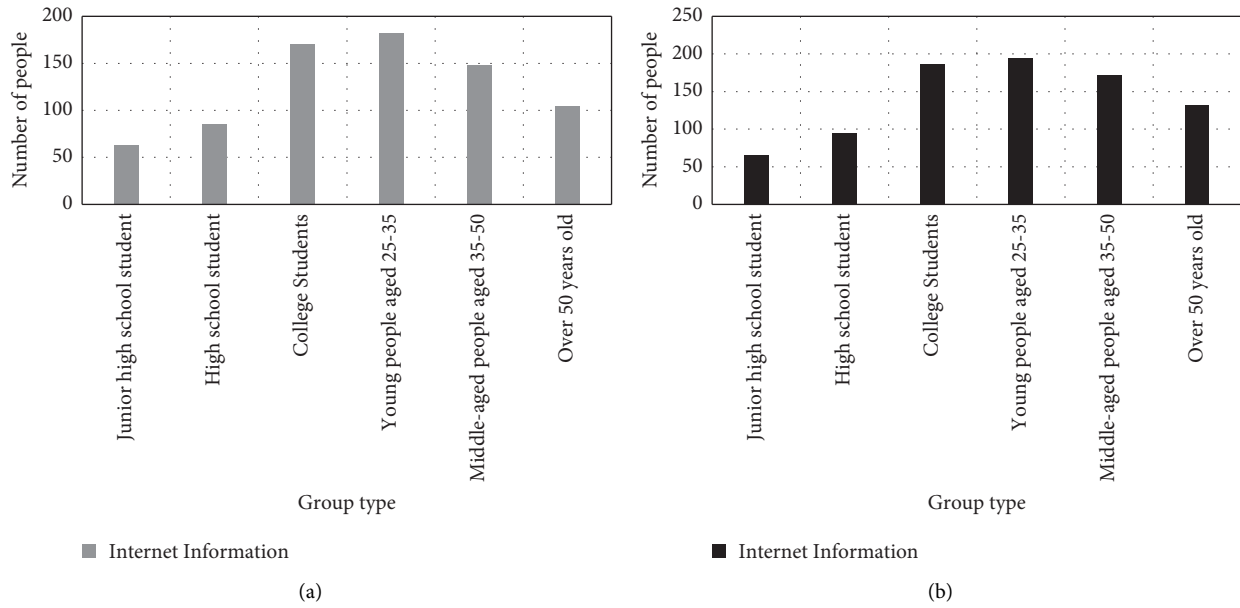
Figure 6: Survey results of network information (a) Men's survey results on Internet information. (b) Women's survey results on online information.

WeChat. After general customers purchase behaviors, they will evaluate their satisfaction with the products, and the customers will display information about the products in the relevant media or describe their own shopping experience. This kind of display activity that occurs on the self-media is generally consumer spontaneous. Therefore, relatively speaking, it is more authentic and credible. Of course, it is also possible for businesses to hire network naval forces to engage in display activities.

Second, the spread of customer online activities on third-party platforms. The proliferation of customer online activities on third-party platforms is reflected in the uncertainty of the space for online activities. It can be said that on the entire Internet platform, customer online activities can exist. The Internet is a "global village," so customer online activities can theoretically spread arbitrarily on the Internet platform. At the same time, different consumers have different interests and hobbies, and the space of online activities diffusion generally conforms to the interests of customers. Because different consumers browse different platforms, the diffusion of customers' online activities has a strong uncertainty. Of course, due to its relative independence, customer online activities appearing on third-party platforms are essentially not under the control of brand marketing, but are also the self-expression of customers themselves. Therefore, in addition to the marketing of the network navy, the authenticity of the information on the third-party platform is relatively high. For example, customers often browse other platforms to learn about them when shopping, or use search engines to view online activities such as related product information of other brands. This article is drawn through a face-to-face questionnaire survey, as shown in Figure 7. 31.6%, 23.85% and 18.56% of its users agree that they will search for product information or product reviews on the Internet. This shows that their

purchasing decisions are largely influenced by the information on these third-party platforms. These online activities can reflect the value of customer online activities on third-party platforms. Therefore, online activities on third-party platforms also have an important impact on online brand sales.

(2) Offline Word of Mouth. In addition to third-party platforms, offline word-of-mouth communication from customers outside the Internet will also affect online brand sales. Of course, because the research perspective of this article is based on the Internet, the article only discusses some offline word-of-mouth activities related to online brands and their impact on online activities. Most customers will comment on this purchase after their online brand buying behavior occurs. Except for the evaluations on the Internet platform, this article believes that most of the customer's evaluations after shopping will be spread offline. This study also found through face-to-face questionnaire surveys that, in addition to online search, introductions by relatives and friends are also an important channel for customers to obtain product or service information. As shown in Figure 8, its proportion is as high as 29.87%.

Obviously, this kind of communication spreads the influence of online brands beyond the Internet, and has a significant impact on online brand sales. For example, after customers purchase a certain brand of goods online, they will spread word of mouth on real social relationships, such as relatives and friends. Of course, most of this kind of word-of-mouth communication obviously appears outside the Internet platform, which has a significant effect on attracting potential target consumers. If this word-of-mouth communication is positive, it will encourage more customers to buy a certain brand online. Conversely, negative offline word-of-mouth will affect online brand reputation.
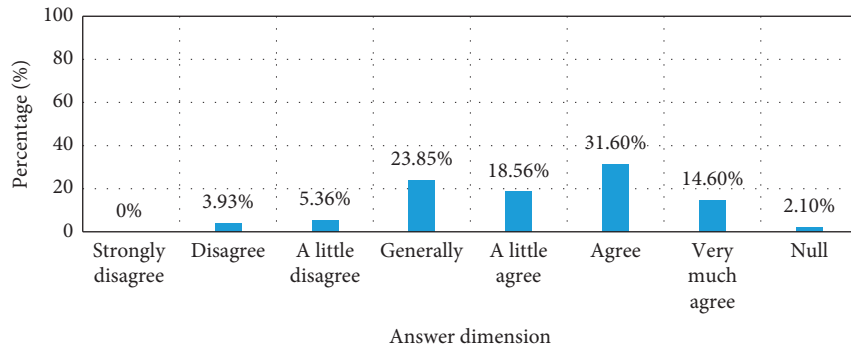
Figure 7: Histogram of product information searched by third-party platforms or search engines.
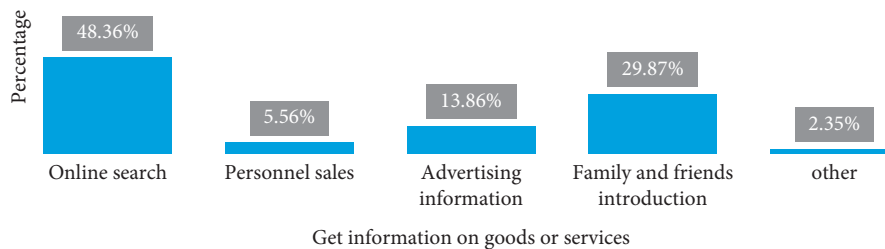
Figure 8: Column chart of channels for online shopping groups to obtain information on goods or services.

If potential customers are attracted by offline word-of-mouth communication and enter the Internet space, the role of offline word-of-mouth communication has become prominent. And new potential customers often have the following two online behaviors:

The first is to seek online word of mouth.

Because the customer obtains the relevant information of the brand offline, the information about the brand has a high degree of uncertainty, and the customer is eager to seek more information to eliminate this uncertainty. Therefore, online activities for seeking brand reputation have emerged. Customers can obtain more relevant information by searching, posting, consulting and other means. Therefore, this kind of customer activity itself promotes the enrichment of online information resources, and can provide relevant information content to latecomers.

The second is to verify offline word-of-mouth behavior.

It is not difficult to see that on the one hand, customers search for online word-of-mouth information through various online methods, and on the other hand, customers will also verify offline word-of-mouth behavior. Because customers cannot be sure of the authenticity of the information provided by offline word-of-mouth, the process of seeking information is also a process of verification. If online word-of-mouth and offline word-of-mouth match, then the customer's brand impression will deepen, thereby forming trust in the brand. On the contrary, if the information that the customer seeks on the Internet platform does not match the information received offline, then the customer's sense of distrust will increase, which will affect the evaluation of the

brand. In summary, whether it is seeking online word-of-mouth behavior or verifying offline word-of-mouth behavior, the customer's activity itself has already constituted an impact on online brand sales. The results of the survey are shown in Figure 9.

The advent of the information age has made it easy for people to obtain various information resources. The survey shows that the majority of people obtain online brand-related information through other channels. As shown in Figure 9, men and women aged 25–35 will search for third-party platforms to learn relevant online brand information, and the number of them is more than 180. People aged 35–50, especially those over 50, prefer to obtain relevant brand product information through recommendations from relatives and friends, and the number of them is more than 130.

It can be seen that the influence of online brand sales comes from five aspects. The article only summarizes the five main factors that affect online brand sales. In the actual Internet environment, there may be other online activities that affect online brand sales. This article can only provide important influencing factors, and cannot completely describe all the influencing factors of online brand sales. In addition, the five aspects will interact and influence each other internally.

4.3. Suggestions on Online Brand Marketing Strategies. As mentioned above, the five factors of online evaluation, brand awareness, online shopping community, online news, and other channels affect online brand sales. Accordingly, this article summarizes the impact of five factors on online brand sales and give some marketing suggestions, which are mainly reflected in the following aspects:
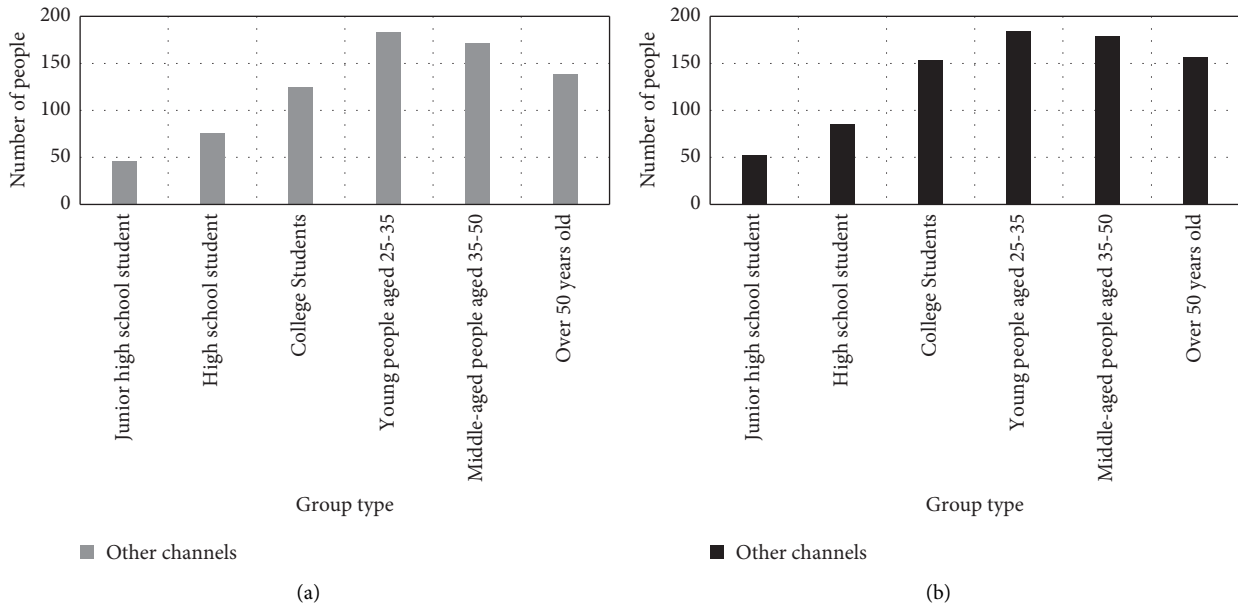
FIGURE 9: Survey results from other channels (a) Men's survey results on other channels (b) Women's survey results of other channels.

First, it enriches potential online information resources.

In the era of new media, information about related products or services can be gathered together through hyperlinks and other means, so that users can obtain massive amounts of information about products or services. If a user finds an advertisement for a certain product or service on the Internet, he will use various Internet tools to understand the product. If the product does not have complete information, customers can search for it online. Therefore, from the perspective of online information content, potential online information resources will continue to be enriched due to the continuous contributions of customers. At the same time, a large number of customer online activities are gathered together in the form of keywords, providing customers with complete online information resources through hyperlinks. Therefore, based on the massive information of Internet products and the means of hyperlinks, customers can learn about all product systems from individual product information, as well as information about products and related brands. Therefore, in the era of new media, with the continuous integration of information, the content of online information resources has been greatly enriched. This has increased online brand sales to a certain extent.

Second, the promotion of online channels.

In the era of traditional media, the interaction between customers and businesses is extremely restricted due to the long time for feedback and limited feedback channels. For example, advertisement information in newspapers cannot provide timely feedback on the product. Even after the feedback, the time for the company to receive the feedback information is very long. Therefore, it is very important to make full use of various feedback channels in the new media era to interact with online customers in an integrated way. With the development of Internet technology, there are more and more ways to conduct product promotion online. For example, after consumers see relevant product information on the Internet, they can immediately click on the product's website or link to view it. At the same time, they will spread relevant product or brand information in real time through various media around them. In the current era of increasingly rich media methods, interactivity has become one of the prominent signs of communication between customers and brands. With the development and popularization of various applications on the Internet, the online activities of customers on various applications are interactive, multi-directional, and participatory, which greatly promotes the promotion of online brands or products. Therefore, the upgrading of online promotion channels is indispensable, such as blog promotion, WeChat promotion, and various application software promotion.

Third, offline word-of-mouth and online word-of-mouth are combined.

In addition to its own online promotion of brand reputation, customers also play a decisive role in the spread of brand reputation. Therefore, it is necessary to combine offline word-of-mouth with online word-of-mouth.

4.3.1. Offline Reputation of Online Brands. The offline reputation of a brand is formed through communication between people. In daily life, customers obtain online promotion information of related brands by communicating with people around them. Out of trust in people around them, they are more convinced of the authenticity and reliability of related brands. Each customer is actually playing the dual role of information disseminator and information receiver. The information that customers get about online brands is more likely to be the experience dissemination of the information disseminator through the experience of purchasing the online brand in person, or it may be the information about the online brand that the information disseminator obtains from other customers. Although there

are differences in the personal needs of each customer, this has also prompted customers to recognize the quality of related brands more. The spread of online brands by customers offline has gradually formed the offline reputation of online brands.

4.3.2. Online Reputation of Online Brands. The online reputation of online brands is mainly formed through online customer reviews. In the era of information sharing, after customers purchase related branded goods online, online platforms can make text comments on related branded goods or services, making relevant information about online brands more intuitive and transparent. Among these reviews, there are not only reviews of related brand products, but also those of shopping platforms, and there are both positive and negative reviews. The spread of these comments through the Internet has gradually formed the online reputation of online brands.

4.3.3. Combining Offline Word-of-Mouth with Online Word-of-Mouth. From the perspective of online brand communication, the development of a brand cannot rely on online or offline unilateral communication. Therefore, the word-of-mouth of an online brand needs to be formed by the combination of online and offline. Customers make reviews when doing purchases, and brands can continuously improve their services based on customer reviews to further meet customer needs. This shows that customers play a leading role in the formation and development of online brands.

Fourth, the communication of the online brand itself is combined with that of customers.

From the perspective of communication, customers are both consumers and communicators. Therefore, the combination of online branding and customer communication will generate a better effect. For example, when customers search for relevant brand product information, they will communicate with brand staff in a timely manner to obtain the information they want. Of course, customers can also consult various experts through various communities and forums, and all this information can only be obtained by searching for them. The propagation process is shown in Figure 10. At the same time, on the Internet platform, all consumer information can be disseminated a second time, and customers can forward it through various platforms such as Weibo, WeChat, and communities. For example, through various self-media means, customers can forward the information content of their favorite products or services to relevant people or recommend them to other friends. This kind of dissemination that customers actively forward will make the secondary dissemination wider to a certain extent. Therefore, customer communication is as important as the brand's own communication, which requires special attention.

Fifth, customers collaborate to participate in online brand communication.

Online brand communication activities require the active participation of customers, so the brand needs to mobilize the enthusiasm and enthusiasm of customers. At the same time, it is also necessary to strengthen the interaction with customers, so as to provide useful information to the brand during the interaction process in order to better carry out brand communication activities. For example, A is a popular search client. When general consumers do not understand brand information, they can search for answers to related brand information through A, and A generally has similar questions or answers from previous consumers, which can meet the needs of consumers. Of course, customers can also open an account A to provide their views on the brand, and it is precisely because of this that makes A's information more comprehensive and true, and provides more comprehensive information for future customers. It can be seen that the collaborative participation of customers in online brand communication has become a norm on the Internet, and the process of participation is also the process of brand information integration. Online brands can also extract the essence of relevant information. Customers can feed back good products or services to the brand, so that relevant brand information can be better displayed to customers, thereby increasing the credibility and promotion of the brand.

5. Discussion

This article conducted 200 effective online questionnaires for six groups of college students and young people aged 25–35 to collect relevant data to analyze the influencing factors of online shopping brand sales. The main findings of this article are analyzed as follows:

(1) Online evaluation. More than 180 men and women aged 25–35 said that they would pay attention to the evaluation of related brand products when shopping;

(2) Brand awareness. Young people aged 25–35 will pay more attention to brand awareness, and college students also express that they prefer branded products with brand awareness, and the number of them is more than 110. Compared with women, men will recognize the quality of branded goods better, and the number of men from college students to men over 50 years old is more than 110;

(3) Online shopping community. The survey results in this article show that young people are more active and proactive in participating in the community. People after 35 years old, because of different interests and low mobile phone usage, platforms such as communities are not the main channels for understanding product information;

(4) Network information. In the era of information sharing, adults will quickly learn about online news through platforms such as TV and have corresponding likes and dislikes. Therefore, online brands need to actively guide the brand's correct and true online news to increase the credibility of the brand to customers.

(5) Channels to learn about other brand products. It mainly focuses on the two aspects of third-party
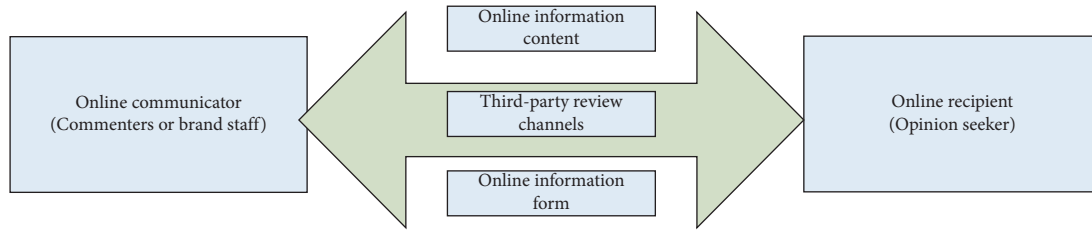
FIGURE 10: Online communication process.

platforms and offline word-of-mouth communication. The 25–35-year-old youth association learns about relevant brand information through the third online platform, and the number of them is more than 180. The majority of men and women aged over 35 learn about relevant product information through the introduction of relatives and friends, and the number is more than 170.

It is not difficult to see that five aspects, including online evaluation and brand awareness, all affect online brand sales.

6. Conclusion and Outlook

Based on the big data processing technology of the Internet of Things, this paper analyzes the influencing factors of online shopping brand sales. It conducts an online questionnaire survey of people from junior high school to over 50 years old. According to the survey, the five aspects of online evaluation, brand awareness, online shopping community, related brand network news, and other brand product information acquisition channels all affect online brand sales. This article puts forward some marketing strategies in a targeted manner. First, by enriching online resources, customers can retrieve and consult relevant brand product information and learn about related services through major platforms. Second, through online promotion, customers can see brand product information on their frequently active platform, and can learn product information by clicking on the link to complete the purchase behavior. Third, through the combination of online and offline to spread brand reputation, customers will share and evaluate on third-party platforms after purchase, and the brand can make corresponding adjustments based on customer reviews to further meet the needs of the market. Fourth, in addition to self-dissemination, we should also pay attention to customer communication. Brand staff can conduct publicity based on answering customer questions, and customers will disseminate relevant information on other Internet platforms. In this way, the promotion of the brand has been further increased. Finally, customers collaborate to participate in online brand communication. Online brands can mobilize the enthusiasm of customers, better promote the brand by understanding customer needs, and increase the credibility of the brand.

The article uses a questionnaire to quantitatively analyze the sales of online shopping brands based on the big data processing of the Internet of Things, but there are still many shortcomings. First, the depth of the article research is not enough, only five aspects of online evaluation, brand awareness, etc. Have been investigated and researched, and no more in-depth research has been conducted on a certain factor; second, the data of the article also has its shortcomings. There are many factors that affect online brand sales. The survey sample only selected five aspects as research data, such as consumer psychology, consumer care elements, and so on. These two aspects are the follow-up work of this article.

Data Availability

All the data used are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] W. U. Zhao-Hui, "Research on the application of internet of Things technology to digital museum construction," *Acta Geoscientia Sinica*, vol. 38, no. 2, pp. 293–298, 2017.

[2] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, "IoT-based big data storage systems in cloud computing: perspectives and challenges," *IEEE Internet of Things Journal*, vol. 4, no. 1, pp. 75–87, 2017.

[3] G. Manogaran, D. Lopez, and N. Chilamkurti, "In-Mapper combiner based MapReduce algorithm for processing of big climate data," *Future Generation Computer Systems*, vol. 86, pp. 433–445, 2018.

[4] S. V. Limkar and R. K. Jha, "A novel method for parallel indexing of real time geospatial big data generated by IoT devices," *Future Generation Computer Systems*, vol. 97, pp. 433–452, 2019.

[5] W. Wu, P. T. P. Quyen, and A. A. A. Rivas, "How e-servicescapes affect customer online shopping intention: the moderating effects of gender and online purchasing experience," *Information Systems and E-Business Management*, vol. 15, no. 3, pp. 689–715, 2017.

[6] R. Ladhari, J. Gonthier, and M. Lajante, "Generation Y and online fashion shopping: orientations and profiles," *Journal of Retailing and Consumer Services*, vol. 48, pp. 113–121, 2019.

[7] I. Abbes, Y. Hallem, and N. Taga, "Second hand shopping and brand loyalty the role of online collaborative redistribution platforms," *Journal of Retailing and Consumer Services*, vol. 52, Article ID 101885, 2020.

[8] M. Swapana and C. Padmavathy, "Relationships among dimensions of online second-hand shopping, satisfaction, and repurchase intention," *International Journal of E-Business Research*, vol. 14, no. 1, pp. 89–102, 2018.

[9] C. Angeli, S. K. Howard, J. Ma, J. Yang, and P. A. Kirschner, "Data mining in educational technology classroom research: can it make a contribution?" *Computers & Education*, vol. 113, pp. 226–242, 2017.

[10] T. Li and L. Long, "Imaging examination and quantitative detection and analysis of gastrointestinal diseases based on data mining technology," *Journal of Medical Systems*, vol. 44, no. 1, pp. 1–15, 2020.

[11] M. Marjani, F. Nasaruddin, A. Gani et al., "Big IoT data analytics: architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, no. 99, pp. 5247–5261, 2017.

[12] D. Zhu, "Deep learning over IoT big data-based ubiquitous parking guidance robot for parking near destination especially hospital," *Personal and Ubiquitous Computing*, vol. 22, no. 5-6, pp. 1–8, 2018.

[13] W. Zhao-Yi, H. Zheng-De, Y. Ping, R. Ting, and L. I. Xin-Hui, "Regularity of wind-dispelling medication prescribed by li dong-yuan: a data mining technology-based study," *Digital Chinese Medicine*, vol. 3, no. 1, pp. 20–33, 2020.

[14] M. Palaniswami, A. S. Rao, D. Kumar, P. Rathore, and S. Rajasegarar, "The role of visual assessment of clusters for big data analysis: from real-world internet of Things," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, no. 4, pp. 45–53, 2020.

[15] M. A. Ahad and R. Biswas, "Request-based, secured and energy-efficient (RBSEE) architecture for handling IoT big data," *Journal of Information Science*, vol. 45, no. 2, pp. 227–238, 2019.

[16] S. Rajendran, O. I. Khalaf, Y. Alotaibi, and S. Alghamdi, "MapReduce-based big data classification model using feature subset selection and hyperparameter tuned deep belief network," *Scientific Reports*, vol. 11, no. 1, Article ID 24138, 2021.

[17] D. Kantoci, E. Džanic, and M. Bogers, "From 'big data' to 'smart data': algorithm for cross-evaluation as a novel method for large-scale survey analysis," *International Journal of Transitions and Innovation Systems*, vol. 6, no. 1, pp. 24–47, 2018.

[18] J. Granat, J. M. Batalla, C. X. Mavromoustakis, and G. Mastorakis, "Big data analytics for event detection in the IoT-multicriteria approach," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4418–4430, 2020.

[19] Z. Xu, G. Zhu, N. Metawa, and Q. Zhou, "Machine learning based customer meta-combination brand equity analysis for marketing behavior evaluation," *Information Processing & Management*, vol. 59, no. 1, p. 102800, 2022.

[20] B. Han, X. Yang, Z. Sun, J. Huang, and J. Su, "OverWatch: A Cross-Plane DDoS Attack Defense Framework with Collaborative Intelligence in SDN," *Security and Communication Networks*, vol. 2018, Article ID 9649643, 15 pages, 2018.

[21] C. Zuo, "Defense of computer network viruses based on data mining technology," *International Journal on Network Security*, vol. 20, no. 4, pp. 805–810, 2018.

[22] J. Xuan, G. Deng, R. Liu, X. Chen, and Y. Zheng, "Analysis of medication data of women with uterine fibroids based on data mining technology," *Journal of Infection and Public Health*, vol. 13, no. 10, pp. 1513–1516, 2020.

[23] A. R. Al-Ali, I. A. Zualkernan, M. Rashid, R. Gupta, and M. AliKarar, "A smart home energy management system using IoT and big data analytics approach," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 4, pp. 426–434, 2018.

[24] Z. Li and X. Yu, "Data mining technology for mechanical engineering computer test system," *Mechanical Systems and Signal Processing*, vol. 141, pp. 106628.1–106628.13, 2020.

[25] C. Xu, L. Xu, Y. Lu, H. Xu, and Z. Zhu, "E-government recommendation algorithm based on probabilistic semantic cluster analysis in combination of improved collaborative filtering in big-data environment of government affairs," *Personal and Ubiquitous Computing*, vol. 23, no. 3-4, pp. 475–485, 2019.

Research Article

Protective Mechanism of Proprotein Convertase Subtilisin-Like Kexin Type 9 Inhibitor on Rats with Middle Cerebral Artery Occlusion-Induced Cerebral Ischemic Infarction

Shengxiong Pu,[1] Chaojun Jia,[1] Zhimin Li,[1] and Yunhua Zang [2]

[1]*Department of Neurology, Affiliated Hospital of North Sichuan Medical College, Nanchong, China*
[2]*Qingdao Hospital of Traditional Chinese Medicine (Qingdao Hiser Hospital), Qingdao, China*

Correspondence should be addressed to Yunhua Zang; yunhua_zang@stu.ahu.edu.cn

The aim of this study was to research the mechanism of proprotein convertase subtilisin-like kexin type 9 (PCSK9) inhibitor in neural protective effect on rat cerebral ischemic reperfusion injury (I/RI). The transient middle cerebral artery occlusion (tMCAO) model of rats was prepared by the suture method, and PCSK9 inhibitor was injected intraperitoneally immediately after I/R. The rats were scored for neurological deficits and the cerebral infarction volume was measured. The brain tissues were collected and western blot (WB) was used to detect the expression of PCSK9. The rat cortical neural stem cells were treated with oxygen glucose deprivation (OGD) to establish a cell model of ischemia/reperfusion. WB was used to detect the expression of PCSK9 and the apoptosis-related pathway proteins. After interfering with the expression of PCSK9 siRNA, the cell viability (cell counting kit-8 assay) and apoptosis (TUNEL staining, Annexin V/PI method) were detected, and the cell proliferation was detected by EdU staining and flow cytometry. The expression of PCSK9 in the brain tissue of the MCAO group was dramatically increased. PCSK9 inhibitor can improve neurobehavioral scores and reduce apoptosis and infarct volume. An OGD model of neural stem cells in vitro was constructed. Inhibiting PCSK9 with si-PCSK9 can increase cell viability, promote cell proliferation, and also reduce cell apoptosis. Inhibition of PCSK9 can decrease the cerebral infarct volume in rats with cerebral I/RI and improve the neural function. Mechanically, inhibition of PCSK9 can lead to the decrease of nerve cell apoptosis and promotion of cell proliferation.

1. Introduction

The current population aging has become a global problem. Correspondingly, ischemic stroke, as one of the main cerebrovascular diseases that threaten human health, has the characteristics of high morbidity, high recurrence, high disability, and high mortality [1]. Because of the existence of collateral circulation in the ischemic penumbra, it is not completely infarcted, and neuronal function is still possible to recover. It has always been a hot spot in the study of acute ischemic stroke. The early restoration of blood flow is of great significance for reducing the infarct size, restoring patient dysfunction, and reducing mortality and disability [2]. After the blood supply of the tissues and organs is interrupted and blood flow is restored, the functions of the tissues and organs are not restored or even worsened, and

the dysfunction and structural damage are more serious, which is called ischemic reperfusion injury (I/RI) [3]. The mechanism of I/RI is very complicated. It is currently believed that various mechanisms such as inflammation, apoptosis, necrosis, oxidative stress, calcium overload, and free radical generation are involved [4]. Thrombolysis is a clinically effective treatment method, but its limited time window limits its clinical application, making the development of new and effective targets and drugs for the treatment of ischemic stroke particularly important.

Proprotein convertase subtilisin-like/kexin type 9 (PCSK9) was discovered by Seidah et al. [5] in 2003, and it is the ninth molecule in the proprotein convertase (PC) family. PCSK9 is mainly produced by the liver and then released into the blood. It returns to the liver through the circulation and binds to the low-density lipoprotein receptor (LDLR) on

the surface of liver cells to form a LDLR-PCSK9 complex, which is digested by hepatocytes [6], leading to a decrease in LDL clearance and an increase in LDL-C in the blood. PCSK9 is mainly expressed in the kidney, small intestine, liver, and other organs. PCSK9 is also expressed in the cortex and hippocampus [7]. A large number of literature studies have pointed out that the mechanism of PCSK9 in the nervous system is different from that of the liver. It is involved in such as inflammation and oxidative stress [8]. In addition, foreign scholars believe that PCSK9 as a neuro apoptosis regulating converting enzyme is also involved in mechanisms such as apoptosis [8, 9].

Pep2-8 is a PCSK9 inhibitor, which can selectively bind to PCSK9 and block various biological effects of PCSK9 [10]. Therefore, we hypothesized that inhibition of PCSK9 expression after I/R may rescue neurons away from ischemia-induced apoptosis via downregulating caspase-3 pathway. In our study, we established a tMCAO model in SD rats to observe the interventional effects of PCSK9 inhibitor on neurobehavioral scores and cerebral infarct volume after ischemia/reperfusion in SD rats. A neural stem cell OGD in vitro was constructed to explore the effects and potential mechanism of PCSK9 on cell proliferation and apoptosis.

2. Methods and Materials

2.1. Experiment Animal.
A total of 30 adult male Sprague-Dawley (SD) rats (body weight 250–280g) and fetal rats at 18 days of gestation were provided by the Experimental Animal Center of North Sichuan Medical College, and adult male SD rats were used for in vivo animal experiments. Fetal SD rats are used for the extraction and culture of neural stem cells in vitro. This study was approved by the Animal Ethics Committee of North Sichuan Medical College Animal Center.

2.2. Animal Grouping and MCAO Model.
SD rats were randomly divided into the following: (1) sham group: rats underwent sham operation. (2) MCAO group: SD rats were reperfused after 2 hours of ischemia. (3) PCSK9i (Pep2-8) treatment group: reperfusion after 2 hours of ischemia, and intraperitoneal injection of $10 \mu g/kg$ PCSK9 inhibitor immediately after reperfusion. The operation method refers to the MCAO method: (1) after the rats were weighed, they were anesthetized by intraperitoneal injection with 4% sodium pentobarbital. (2) The anesthetized rat was fixed on the operating table in a supine position. After the neck was shaved and disinfected, an incision with a length of about 2 cm was cut in the middle. (3) The blood vessels and the vagus nerve were separated, and the left common carotid artery, external carotid artery, and internal carotid artery were exposed. (4) The proximal end of the external carotid artery and the common carotid artery were ligated, a single knot was tied at the distal end of the common carotid artery, and the internal carotid artery was clamped with an arterial clip. (5) A small gap was cut midway above the ligation of the proximal end of the common carotid artery, and the prepared nylon suture was slowly advanced through the

incision to the internal carotid artery. At the same time, the arterial clamp on the internal carotid artery was loosened. (6) When the length of the thread plug was 18 mm–20 mm, the resistance can be obviously felt. At this time, the opening of the middle cerebral artery was blocked and the thread plug was fixed. (7) After blocking for 2 hours, the single knot was loosened, the nylon thread plug was slowly pulled out, and the distal end of the common carotid artery was ligated, finally restoring blood circulation and achieving reperfusion. (8) After suturing the skin, debridement and disinfection were performed. Sham group: arterial embolization by inserting a thread plug was not performed, and the other operations were the same as those in the MCAO group.

2.3. Modified Neurological Severity Scores.
mNSS is one of the most commonly used scales for evaluating neurological function after stroke, including four categories: motor, sensory, reflex, and balance. The rat was assigned a value when it cannot complete the action or meets the item. The greater the score, the more severe the damage. The scoring system referred to the method in the previous literature [11].

2.4. Rotarod Test.
The Rotarod test was mainly used to evaluate the coordination and balance of rodents. The rat was placed on a roller, the speed was gradually increased from 4 rpm to 40 rpm within 5 minutes, and the incubation period was recorded from the beginning to the time of the rat falling. The rats were trained every day for three days before modeling, and the average was measured three times.

2.5. Frozen Section of Mouse Brain.
On the 7th day after modeling, the rats were anesthetized with 4% sodium pentobarbital. After opening the chest, the right atrial appendage was cut and slowly 50 ml of normal saline was infused through the left ventricle until the right atrial appendage was clear. Then, the brain tissue was taken out and quickly put in isopentane precooled in the ultralow temperature refrigerator for about 1-2 minutes. Then, the rat brain was taken out, marked, and put in the ultralow temperature refrigerator for later use. When slicing, the rat brain was taken out and placed in a microtome for about 1 hour. After adaptation, the slice was embedded in OCT (Keygen, Nanjing, China) with a thickness of $20 \mu m$.

2.6. TTC Staining.
The brain tissue sections were incubated with 2% TTC solution (Jian Cheng, Nanjing, China) at 37°C for 20 minutes, and then the excess TTC regent was drained. The brain slices were fixed with 4% paraformaldehyde. The Image J analysis system was used to quantify the infarct volume.

2.7. Western Blot.
After the brain tissue or neural stem cells total protein was extracted and determined, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and membrane transfer were performed. After blocking, the primary antibody (PCSK9, Abcam, Cambridge, MA, USA,

mouse, 1:1000; Bax, Abcam, rabbit, 1:1000; Bcl-2, Abcam, rabbit, 1:1000; caspase-3, Abcam, rabbit, 1:1000; and β-actin, Abcam, 1:5000) was added and incubated overnight at 4 °C. The secondary antibody conjugated with horseradish peroxidase was incubated for 1 hour at room temperature, and the membrane was washed 3 times with tris buffered saline-tween (TBST). Then,the membrane was incubated with electrochemiluminescence (ECL) solution (Camilo, Nanjing, China) for 1 minute and exposed for 5–20 minutes in the cassette for developing and fixing. AlphaEaseFC grayscale analysis software was used to analyze the band grayscale scanning, and the results were expressed as the target protein by the gray value of the sample band/corresponding internal reference gray value. In this study, β-actin was used as an internal reference.

2.8. Isolation of Primary Cortical Neural Stem Cells. Primary neural stem cells were taken from the cerebral cortex of rat embryos at 18 days of pregnancy. The cerebral cortex was digested with 0.25% trypsin (Camilo, Nanjing, China), and then the cell suspension was inoculated into a 6-well plate under 95% air and 5% carbon dioxide conditions. The cell culture medium was DMEM containing 10% fetal bovine serum (FBS) and cytosine-D-arabinofuranoside ($10 \mu M$). In OGD treatment, cortical neurons were cultured in Dulbecco's modification of Eagle's medium (DMEM, Thermo Fisher Scientific, Waltham, MA, USA) for 12 hours under normal conditions and then incubated for 2 hours under hypoxic conditions (95% N_2 and 5% CO_2) with 0.5 mmol/L of glucose and sodium dithionite, and the medium was changed every two days and cultured for 7 days.

2.9. Transfection. The neural stem cells in the logarithmic growth phase were seeded in a 96-well plate to grow the cell density to 60%~80%, and the cells were transfected according to the Lipofectamin^{2000} (Keygen, Nanjing, China) instructions. siRNA-PCSK9 and negative control were diluted with $250 \mu L$ serum-free medium and incubated at room temperature for 5 minutes. Lipofectamin^{2000} $5 \mu L$ was diluted with $250 \mu L$ opti-MEM without serum and incubated at room temperature for 5 minutes. Then, the above liquid was mixed, incubated at room temperature for 20 minutes, and then added to the cell culture well. Cells were incubated in a 37 °C, 5% CO_2 incubator for 6–8 hours. After culturing for 24–48 hours, subsequent experiments were performed.

2.10. Cell Counting Kit-8 (CCK-8) Assay. The treated neural stem cells (2×10^4 cells/$100 \mu L$) were inoculated for 96 hours and $10 \mu L$ of CCK-8 solution (Ye Sen, Shanghai, China) was added to each sample, 3 replicate wells were repeated. After incubating for 2–4 hours, the absorbance at 450 nm was measured with a microplate reader (BD-Biosciences Franklin Lakes, NJ, USA).

2.11. Cell Cycle Assay. The treated neural stem cells were cultured in a 6-well culture plate at 37 °C until the cells reached 80% confluence. The cells were collected with 70% precooled ethanol. After washing with phosphate buffered saline (PBS) twice, the cells were resuspended with 0.5 mL PBS and added with PI (Thermo Fisher Scientific, Waltham, MA, USA) for 30 minutes. Finally, the cells were analyzed by flow cytometry (BD-Biosciences, USA).

2.12. Flow Cytometry. The processed neural stem cells were collected first, and then the cells were suspended in a buffer containing $60 \mu g/ml$ PI and Annexin V (Ye Sen, Shanghai, China) and kept at room temperature in the dark for 15 minutes.Subsequently, cell apoptosis was detected by ACS-Caliber flow cytometer.

2.13. TUNEL Staining. After the sections or cell slides were incubated with 2% hydrogen peroxide for 30 minutes, the cell membrane was permeated with Trixon X-100 (Keygen, Nanjing, China) for 30 minutes. Then, the TdT enzyme buffer (Keygen, Nanjing, China) was used to incubate at room temperature for 15 minutes, and then $50 \mu L$ of TdT enzyme reaction solution was added dropwise and incubated at room temperature for 60 minutes. After washing with PBS, the reaction was terminated. Next, the peroxidase-labeled antibody was used for reaction at room temperature for 30 minutes. After 10 minutes of methyl green counterstaining, 4 fields of view were randomly selected with an image analysis system to count the number of positive apoptotic cells.

2.14. 5-Ethynyl-2'-Deoxyuridine (EdU) Staining. Neural stem cells treated with OGD were seeded in a 6-well plate at 1×10^5 cells per well. After culturing overnight, the original medium was replaced with a fresh medium containing EdU ($100 \mu M$, Camilo, Nanjing, China) for 2 hours, then the cells were fixed with acetone for 20 minutes. The Apollo reaction buffer with FITC fluorescein was incubated for 1 hour and then photographed with a laser confocal fluorescence microscope (Leica, Wetzlar, Germany).

2.15. Statistical Analysis. All data were expressed as the mean ± standard deviation ($\overline{X} \pm SD$) of three repeated experiments and analyzed by Statistical Product and Service Solutions (SPSS) 20.0 statistical software (IBM, Armonk, NY, USA). Student's t-test was used for comparison between the two groups. One-or two-way analysis of variance (ANOVA) was used for data analysis among multiple groups. The statistical analysis of behavioral data adopts repeated measure ANOVA. $p < 0.05$ is regarded as a statistically significant difference.

3. Results

3.1. Expression of PCSK9 Increases in MCAO Rats' Brain Tissue. In order to verify the success of MCAO model, we performed neurobehavioral scores on the rats after modeling. It was found that the mNSS score of the rats increased in the MCAO group rats (Figure 1(a)), and the incubation period of the rat Rotarod test decreased (Figure 1(b)). On the
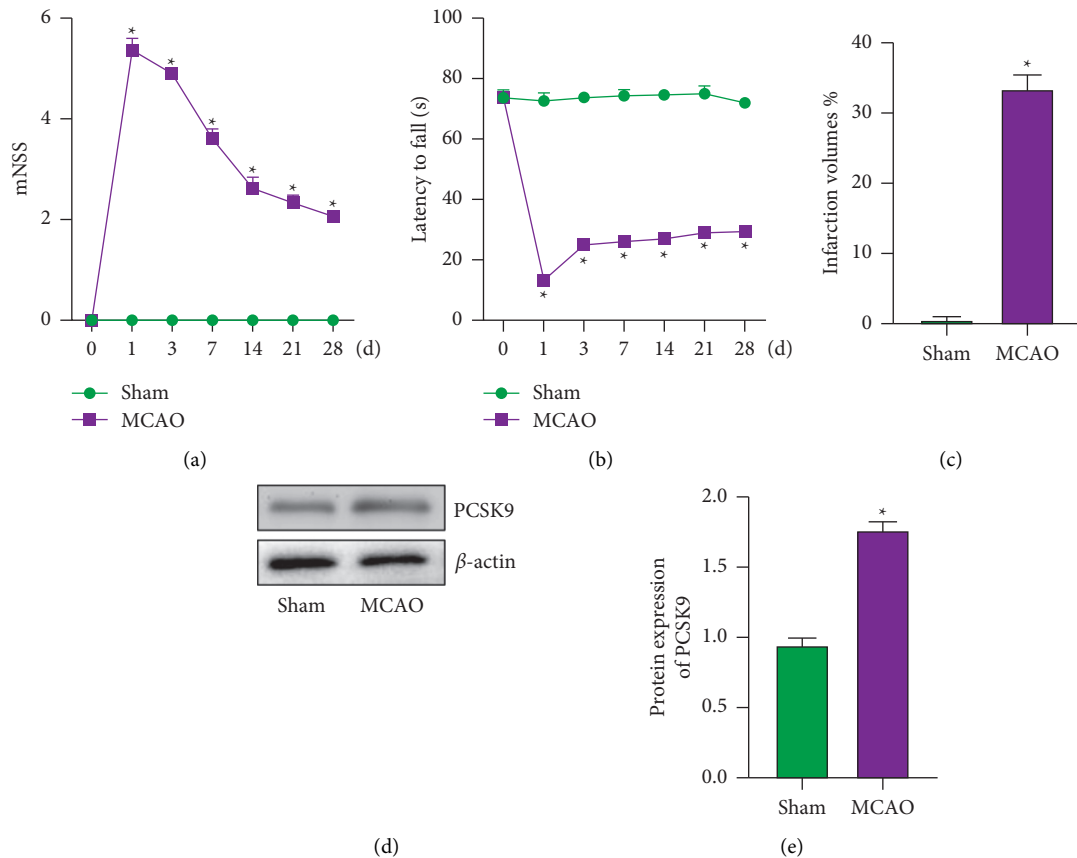
FIGURE 1: The expression of PCSK9 increases in MCAO rats' brain tissue. (a) Modified Neurological Severity Scores (mNSS). (b) Rotarod test. (c) TTC staining detected the infarct volume of brain tissue. (d, e) WB detected the protein of PCSK9 in brain tissue. * *vs.* sham group <0.05.

7th day of I/R, the brains of the two groups (sham/MCAO) were taken for TTC staining. It was found that the cerebral infarct volume of the rats in the MCAO group increased dramatically than the sham group (Figure 1(c)). The above results show that MCAO model was successful. Next, we used WB method to detect the expression of PCSK9 in the subventricular zone of rats. Compared with the sham group, the expression of PCSK9 in the MCAO group was dramatically increased (Figures 1(d) and 1(e)).

3.2. Protective Effect of PCSK9 Inhibitor on MCAO Rats.
In order to explore the effect of PCSK9 on MCAO rats, we injected PCSK9 inhibitor into the intraperitoneal cavity. WB results showed that PCSK9 inhibitor can dramatically inhibit the expression of PCSK9 in brain tissue (Figures 2(a) and 2(b)). At the same time, we performed neurobehavioral scores on rats and found that PCSK9 inhibitor can reduce the score of mNSS and increase the incubation period of the Rotarod test (Figures 2(c) and 2(d)). TTC results showed that compared with the MCAO group, the infarct volume in the MCAO + PCSK9i group was dramatically reduced (Figure 2(e)). Next, we used TUNEL staining to detect apoptotic cells in rat brain tissue and found that compared with MCAO group, the percentage of apoptotic cells in MCAO + PCSK9i group decreased (Figures 2(f) and 2(g)).

The above results indicated that PCSK9i can improve the neurological function of MCAO rats and reduce the volume of cerebral infarction and cell apoptosis.

3.3. Inhibition of PCSK9 Partially Inhibits OGD-Induced Apoptosis via Bax/Bc-l2 Signaling In Vitro.
We used 18-day-old fetal rat cortical neural stem cells to establish an in vitro cell model of ischemia and hypoxia. To explore the effect of PCSK9 on OGD-induced in vitro ischemic-hypoxic cell models, siRNA was used to knock down PCSK9. WB was used to detect the efficiency of knockdown, the results showed that the expression of PCSK9 in the OGD + si-PCSK9 group was dramatically reduced compared with OGD + si-NC group (Figures 3(a) and 3(b)). Furthermore, the apoptosis-related pathway Bax/Bcl-2 signaling was determined by WB assay, showing that OGD treatment significantly increased the expression of Bax and caspase-3 and decreased the Bcl-2 expression, but inhibition of PCSK9 markedly reduced the ratio of Bax/Bcl-2 and the level of caspase-3 (Figures 3(c) and 3(d)). CCK-8 assay was used to detect cell viability and it was found that compared with the OGD group, the cell viability of the OGD + si-PCSK9 group was improved (Figure 3(e)). The result of TUNEL staining showed that the ratio of TUNEL positive cells in the si-PCSK9 group decreased greatly (Figure 3(f) and 3(g)), and

(a)

(b)

(c)

(d)

(e)

(f)

(g)
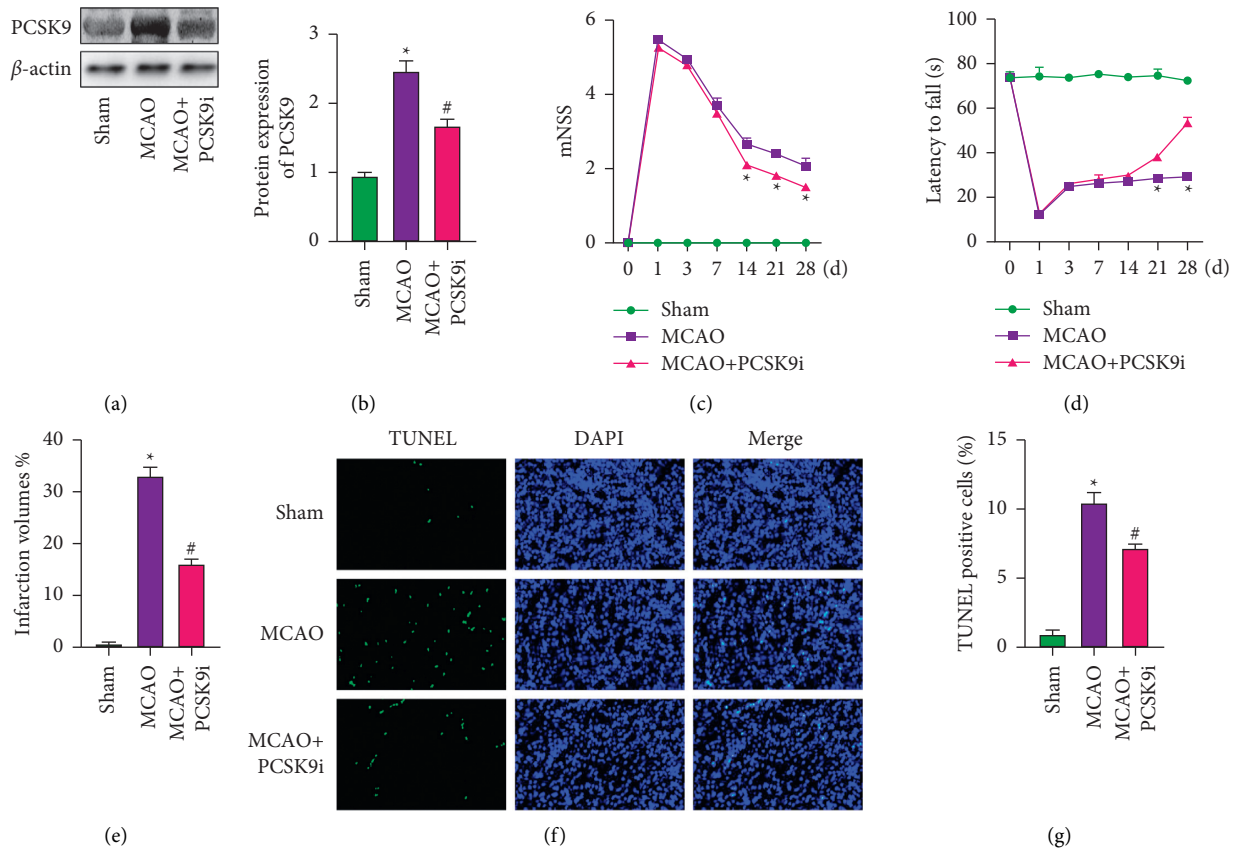
FIGURE 2: Protective effect of PCSK9 inhibitor on MCAO rats. (a, b) WB detected the protein of PCSK9 in brain tissue. (c) Modified Neurological Severity Scores (mNSS). (d) Rotarod test. (e) TTC staining detected the infarct volume of brain tissue. (f, g) TUNEL staining detected the apoptosis cell of brain tissue. * *vs.* sham group <0.05; # *vs.* MCAO group <0.05.

flow cytometry results showed that si-PCSK9 can dramatically reduce the rate of apoptosis (Figures 3(h) and 3(i)). The above results indicated that inhibiting PCSK9 partially inhibits OGD-induced cell apoptosis.

3.4. Inhibiting PCSK9 Promotes the Proliferation of Neural Stem Cells. EdU staining test results showed that compared with the control group, cell proliferation in the OGD group was dramatically inhibited. Compared with the OGD + si-NC group, si-PCSK9 alleviated the cell proliferation inhibition caused by OGD (Figures 4(a) and 4(b)). In the cell cycle analysis, compared with the control group, the number of cells in the OGD group was dramatically upregulated in the G0/G1 phase and dramatically downregulated in the S phase. Compared with the OGD + si-NC group, the number of cells in the OGD + si-PCSK9 group was dramatically downregulated in G0/G1 phase and dramatically upregulated in S phase (Figure 4(c)). All the results showed that inhibiting PCSK9 can promote the proliferation of neural stem cells.

4. Discussion

In this study, we first used the rat MCAO model to find that the expression of PCSK9 in brain tissue was increased. In in vivo experiments, we found PCSK9i can improve neurobehavioral

functions and reduce apoptosis and infarct volume. At the same time, neural stem cells were used to establish an OGD model, and si-PCSK9 was used to inhibit the expression of PCSK9 to clarify the role of PCSK9 in the OGD model in vitro. Our results showed that after OGD treatment, the expression of PCSK9 was dramatically increased, and si-PCSK9 inhibited OGD-induced apoptosis and promoted proliferation (Figure 5).

The mechanism of cerebral ischemia/reperfusion injury is very complicated, including inflammation, apoptosis, and mitochondrial damage. PCSK9 is a serine protease related to the regulation of blood lipids discovered in recent years [12]. It is mainly produced by the liver and then released into the blood to regulate LDL in the blood. However, recent studies have found that PCSK9 is also expressed in the cerebral cortex and hippocampus of mice. After making a mouse ischemia/reperfusion model, Rousselet et al. [13] detected that PCSK9 was dramatically upregulated. In addition, Tang et al. [14] found that PCSK9 can downregulate LDLR levels after an ischemic stroke in mice, leading to hyperlipidemia. Chiang et al. [15] also found that PCSK9 was upregulated after inducing neuronal apoptosis. The above series of experiments show that PCSK9 expression will increase after the onset of ischemic stroke, and some literatures pointed out that the role of PCSK9 in the brain and the liver may be different. In addition to participating in the regulation of

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)
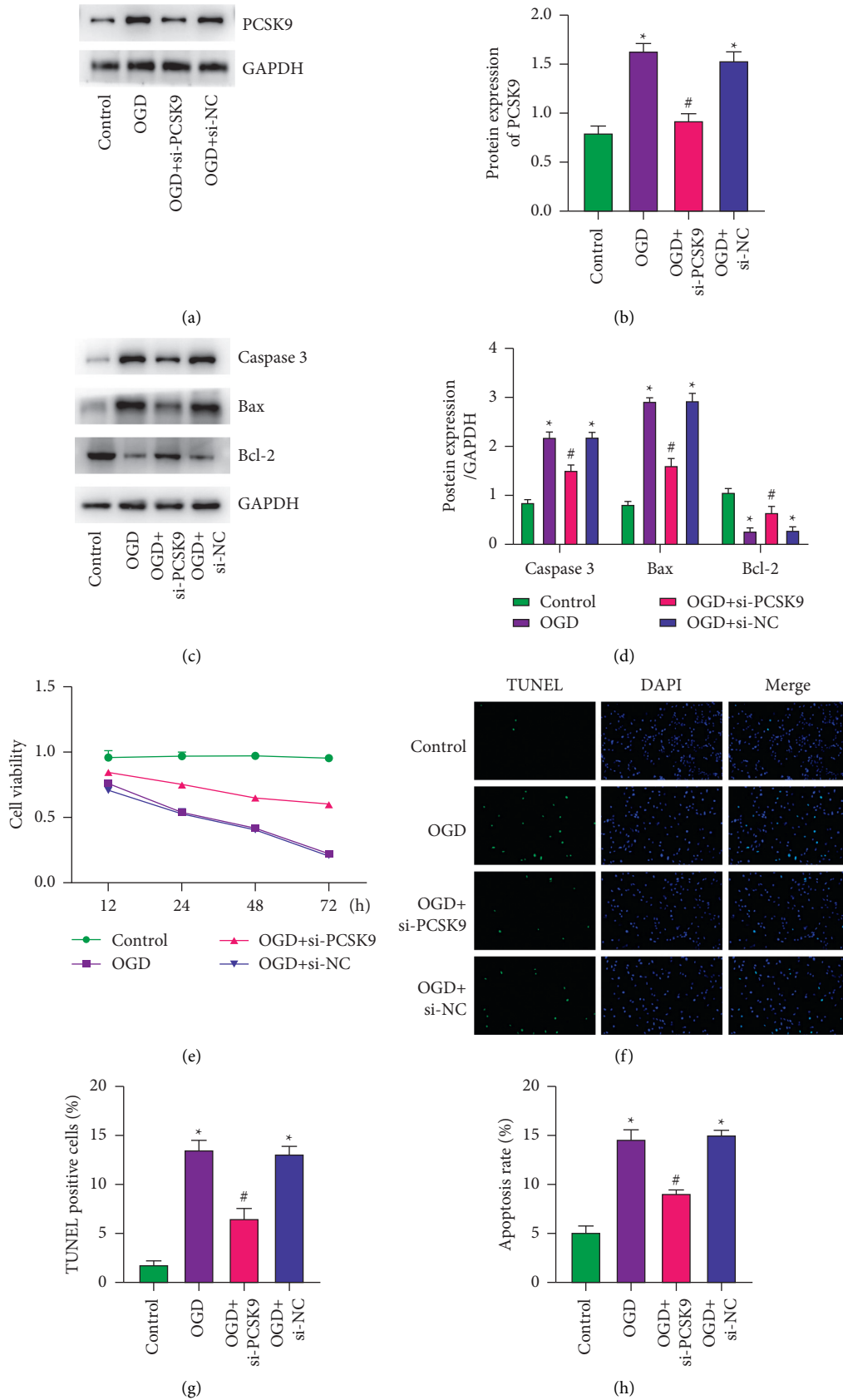
Figure 3: Continued.

(i)
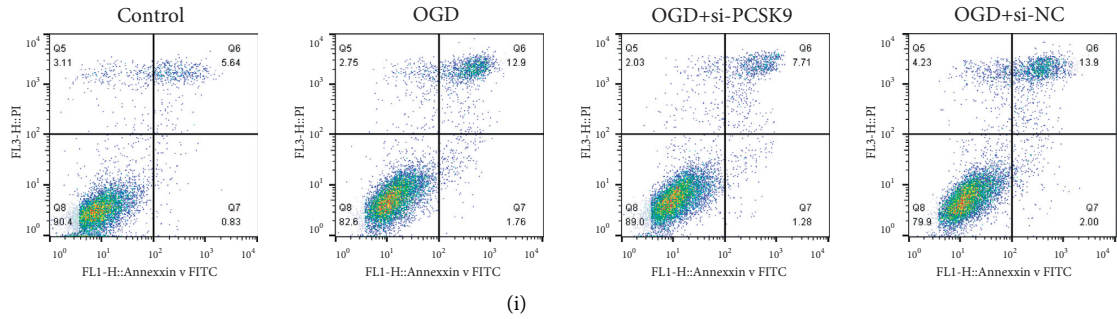
FIGURE 3: Inhibition of PCSK9 partially inhibits OGD-induced apoptosis in vitro. (a, b) WB detected the protein of PCSK9 in neural stem cells. (c, d) WB detected the protein of Bax, Bcl-2, and caspase-3 in neural stem cells. (e) CCK-8 assay detected the cell viability. (f, g) TUNEL staining detected the apoptosis of neural stem cells. (h, i) Flow cytometry detected the apoptosis rate of neural stem cells. * *vs.* control group <0.05; # *vs.* OGD + si-NC group <0.05.

(a)   (b)   (c)
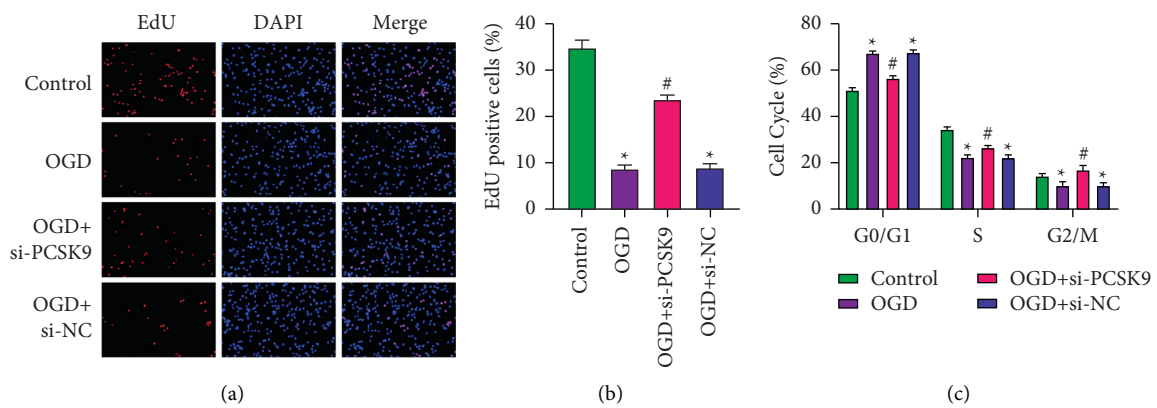
FIGURE 4: Inhibiting PCSK9 promotes the proliferation of neural stem cells. (a, b) EdU staining detected the proliferation of neural stem cells. (c) Flow cytometry detected the cell cycle of neural stem cells. * *vs.* control group <0.05; # *vs.* OGD + si-NC group <0.05.
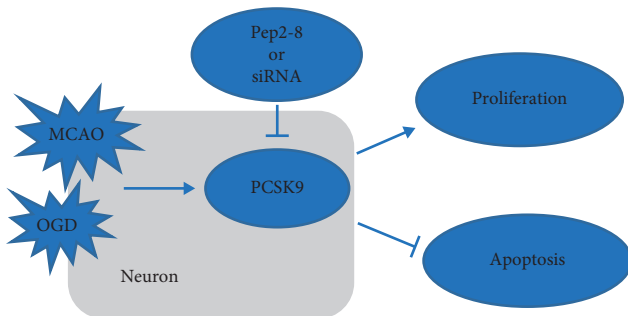
FIGURE 5: Schematic diagram of the protective mechanism of PCSK9 inhibitor on rats with middle cerebral artery occlusion-induced cerebral ischemic infarction.

blood lipids, it also has a wider range of biological effects. It is also involved in insulin resistance, inflammation, cell development, and apoptosis.

Previous studies have found that the MAPK pathway is involved in mediating PCSK9-mediated cell apoptosis [16, 17]. Through the action of shRNA, a study found that PCSK9 dramatically inhibited the phosphorylation of p-p38MAPK and JNK, [18] thereby affecting cell apoptosis. Therefore, we can guess that PCSK9 can directly induce

p-p38MAPK upregulation through the MAPK signaling pathway, thereby mediating neuronal apoptosis. Our study found that si-PCSK9 inhibits OGD-induced cell cycle arrest in G1 phase, indicating that knockdown of PCSK9 may restore the normal progression of the cell cycle through its function as a spindle checkpoint in the cell cycle. However, there are some limitations in the present study. First, we only proved that inhibition of PCSK9 displayed a well correlation with I/R-induced neuronal apoptosis in vivo and in vitro, but the underlying mechanism concerning PCSK9 inhibition regulated neuronal apoptosis was hardly reported. Second, we just preliminarily evaluated the effect of PCSK9 inhibition on function and injured area in MCAO rat model, and more systematic analysis and detection are needed to assess the efficacy in in vivo MCAO rats. Thus, our further researches would aim to the in-depth investigation on the molecular mechanism and therapeutic evaluation of PCSK9 inhibition in ischemic stroke model in the future.

5. Conclusion

Inhibition of PCSK9 can improve the cerebral infarct volume in rats with cerebral ischemia/reperfusion injury,

improve the dysfunction caused by ischemia, and protect the nerve, and its mechanism may be related to the inhibition of nerve cell apoptosis and promotion of its proliferation.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Barthels and H. Das, "Current advances in ischemic stroke research and therapies," *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1866, no. 4, Article ID 165260, 2020.

[2] Y. Song and B. Wang, "MiR-16 exacerbates neuronal cell growth and inhibits cell apoptosis by targeting AKT3 in cerebral ischemia injury," *Tropical Journal of Pharmaceutical Research*, vol. 20, no. 10, pp. 2049–2054, 2021.

[3] S. B Ko and B. W Yoon, "Blood pressure management for acute ischemic and hemorrhagic stroke: the evidence," *Seminars in Respiratory and Critical Care Medicine*, vol. 38, no. 6, pp. 718–725, 2017.

[4] M. Liu, J. Xie, X. Zheng et al., "Correlation between stromal cell-derived factor 1 and the prognosis of transient ischemic attack," *Iran Red Crescent Me*, vol. 23, no. 5, p. e348, 2021.

[5] C. Jung, G. G. Tomás, J. Selent et al., "A gain-of-function SNP in TRPC4 cation channel protects against myocardial infarction," *Cardiovascular Research*, vol. 91, no. 3, pp. 465–471, 2011.

[6] R Schulz and K. D Schlüter, "PCSK9 targets important for lipid metabolism," *Clinical research in cardiology supplements*, vol. 12, no. Suppl 1, pp. 2–11, 2017.

[7] L Wang, Z Wang, J Shi et al., "Inhibition of proprotein convertase subtilisin/kexin type 9 attenuates neuronal apoptosis following focal cerebral ischemia via apolipoprotein E receptor 2 downregulation in hyperlipidemic mice," *International Journal of Molecular Medicine*, vol. 42, no. 4, pp. 2098–2106, 2018.

[8] J. Ji, M. Feng, X. Niu, X. Zhang, and Y. Wang, "Liraglutide blocks the proliferation, migration and phenotypic switching of Homocysteine (Hcy)-induced vascular smooth muscle cells (VSMCs) by suppressing proprotein convertase subtilisin kexin9 (PCSK9)/low-density lipoprotein receptor (LDLR)," *Bioengineered*, vol. 12, no. 1, pp. 8057–8066, 2021.

[9] X.-S. Zhao, Q. Wu, J. Peng et al., "Hyperlipidemia-induced apoptosis of hippocampal neurons in apoE(−/−) mice may be associated with increased PCSK9 expression," *Molecular Medicine Reports*, vol. 15, no. 2, pp. 712–718, 2017.

[10] Z. Ding, X. Wang, S. Liu et al., "PCSK9 expression in the ischaemic heart and its relationship to infarct size, cardiac function, and development of autophagy," *Cardiovascular Research*, vol. 114, no. 13, pp. 1738–1751, 2018.

[11] K. L Schaar, M. M Brenneman, and S. I Savitz, "Functional assessments in the rodent stroke model," *Experimental & Translational Stroke Medicine*, vol. 2, no. 1, p. 13, 2010.

[12] A. Helgadottir, S. Gretarsdottir, G. Thorleifsson et al., "Variants with large effects on blood lipids and the role of cholesterol and triglycerides in coronary disease," *Nature Genetics*, vol. 48, no. 6, pp. 634–639, 2016.

[13] E. Rousselet, J. Marcinkiewicz, J. Kriz et al., "PCSK9 reduces the protein levels of the LDL receptor in mouse brain during development and after ischemic stroke," *Journal of Lipid Research*, vol. 52, no. 7, pp. 1383–1391, 2011.

[14] Z. Tang, L. Jiang, J. Peng et al., "PCSK9 siRNA suppresses the inflammatory response induced by oxLDL through inhibition of NF-κB activation in THP-1-derived macrophages," *International Journal of Molecular Medicine*, vol. 30, no. 4, pp. 931–938, 2012.

[15] L. W. Chiang, J. M. Grenier, L. Ettwiller et al., "An orchestrated gene expression component of neuronal programmed cell death revealed by cDNA array analysis," *Proceedings of the National Academy of Sciences*, vol. 98, no. 5, pp. 2814–2819, 2001.

[16] Z. Ding, S. Liu, X. Wang et al., "Cross-talk between PCSK9 and damaged mtDNA in vascular smooth muscle cells: role in apoptosis," *Antioxidants and Redox Signaling*, vol. 25, no. 18, pp. 997–1008, 2016.

[17] J. Li, X. Liang, Y. Wang, Z. Xu, and G. Li, "Investigation of highly expressed PCSK9 in atherosclerotic plaques and ox-LDL-induced endothelial cell apoptosis," *Molecular Medicine Reports*, vol. 16, no. 2, pp. 1817–1825, 2017.

[18] M. Wicinski, J. Zak, B. Malinowski, G. Popek, and G. Grzesk, "PCSK9 signaling pathways and their potential importance in clinical practice," *The EPMA Journal*, vol. 8, no. 4, pp. 391–402, 2017.

Research Article

Risk Assessment and Regulation Algorithm for Financial Technology Platforms in Smart City

Xiaoxi Liu (ID),[1] **Xiaoling Yuan,**[1] **Rui Zhang,**[2] **and Nan Ye**[2]

[1]*School of Economics and Finance, Xi'an Jiaotong University, Xi'an City 710049, Shaanxi, China*
[2]*Xi'an Institute of Space Radio Technology, Xi'an City 710100, Shaanxi, China*

Correspondence should be addressed to Xiaoxi Liu; liuxiaoxi@stu.xjtu.edu.cn

The informatization of cities has been further promoted, and the construction of smart cities supported by technological innovation has been upgraded. The financial industry and data are closely related. Whether the financial industry can make good use of new information technology is the key to its successful transformation. The development of smart cities has a significant effect on the development of people's livelihood, the process of urbanization, the use of technology, the solution of urban problems, and the improvement of economic levels. This also provides a good choice for the development of cities in each country. For better development, it needs technical support. Therefore, it is very important to improve the technical level. This research mainly discusses the risk assessment and regulation algorithms of financial technology platforms in smart cities. This study divides the risk decision channels into two paths based on the smart city theory, considers the internal risk factors and external risk factors of the robo-advisory service platform from the three perspectives of platform characteristics, corporate characteristics, and investor characteristics and exploring the construction of a robo-advisory service platform risk prediction model based on the machine learning perspective. The design and implementation of a personalized financial investment prototype system, a Python-based web development framework Django, and a variety of toolkits have realized a B/S architecture robo-advisor. Among them, the function of buying and selling ETF and the trend recording function after buying are realized by simulating the transaction data collected by the data collection module. The study found that the key potential characteristics that constitute platform risks are mainly the listing year of the background company, the age of the platform, the investment threshold, and the search index. To a certain extent, this provides data support for investors and regulatory authorities to evaluate platform capabilities and platform selection. Investors should comprehensively consider platform qualifications when making platform decisions and pay attention to information such as the age of listing of companies with platform background, platform age, and investment thresholds. Only when the quality of people is improved, the quality of the population of this city improves, so that the development of the city has a broad room for growth. The accuracy of the similar formula calculation method in the big data proposed in this study reached 88%. This research provides new ideas for perfecting the black box regulatory system of robo-advisory algorithms.

1. Introduction

Based on the integration and utilization of resources, smart cities cooperate with terminal equipment through informatization to achieve the improvement of the level of urban governance, thereby enhancing the intelligence and technology of infrastructure. The acceleration of the network upgrade has further promoted the development of various fields. Financial information is constantly emerging, and a large amount of information can indeed help people understand the financial market more fully, but it also increases the difficulty for users to obtain effective information and reduces efficiency. Through the intelligent integration and analysis of data resources, it provides an important reference for government decision-making and has a significant role in urban governance, economic growth, and improvement of people's livelihood.

Internet finance can be regarded as the basic version of financial technology, and financial technology is its upgraded version. Compared with Internet finance, the

penetration and influence of technology elements on financial technology is more obvious. Financial development can accelerate the informatization development of smart cities through technological innovation, and the efficiency of financial development plays the most prominent role. In the long run, the coordinated and orderly development of the overall financial mechanism and the close integration of technology and finance play an important role in improving the level of informatization development and accelerating the pace of smart city construction [1]. On this basis, corresponding countermeasures and suggestions are put forward to promote the integration of financial development and technological innovation to accelerate the pace of construction of smart cities in various regions.

Financial and technological developments promote each other, and financial service trade and technological advancement cherish each other. From a historical point of view, the financial service trade has emerged on the basis of trade in goods and has developed rapidly under the influence of technological progress, regulatory reforms, and economic growth. As one of the traditional industries, the financial industry is closely related to data. Whether the financial industry can make good use of new information technology is the key to its successful transformation. Menouar believes that there is no smart city without a reliable and efficient transportation system, this necessity makes ITS a key component of any smart city concept. Although traditional ITS technologies have been deployed in smart cities globally, the implementation of next-generation ITS relies on the effective integration of connected cars and autonomous vehicles, and these two technologies are undergoing extensive field testing in many cities around the world. Although these two emerging technologies are critical to the realization of fully automated transportation systems, there is still a great need for the automation of other roads and transportation components. Due to its mobility, autonomous operation, and communication/processing capabilities, UAVs are envisioned for many ITS applications. Menouar et al. described possible ITS applications that can use drones and emphasized the potential and challenges of ITS that support drones in the next generation of smart cities [2]. Bo et al. believed that data-intensive analysis is the main challenge facing smart cities due to the ubiquity of various sensors. The natural nature of geographic distribution requires new computing paradigms to provide location-aware and delay-sensitive monitoring and intelligent control. Fog computing, which extends computing to the edge of the network, fits this need. In order to protect the future community, it is necessary to integrate intelligence in the fog computing architecture, for example, perform data representation and feature extraction, identify abnormal and dangerous events, and provide the best response and control. They used an intelligent pipeline monitoring system based on fiber optic sensors and sequential learning algorithms to analyze case studies to detect events that threaten pipeline safety. A working prototype was built to experimentally evaluate the event detection performance of identifying 12 different events. These experimental results prove the feasibility of the system to be implemented in the city in the future [3]. Brundu et al. introduced an IoT software infrastructure that supports new control policies for energy management and simulated urban areas. The platform Brundu et al. proposed enables (near) real-time building energy profile interoperability and correlation with environmental data from sensors and building and grid models. In the smart city environment, the platform realizes the integration of heterogeneous data sources at the building and regional levels, and the simulation of new energy policies at the regional level. It aims to optimize energy use and consider its impact. The platform Brundu et al. proposed has been deployed in a real area, and a novel heating network control strategy has been developed and tested [4]. Mosannenzadeh believed that smart energy cities are an emerging concept in urban development, aiming to optimize the urban energy system and improve the quality of life of citizens. However, the development of smart energy cities requires a well-defined and consistent conceptual core to ensure its accurate interpretation and successful implementation. He defined the development of smart energy cities not only in the theoretical context but also in terms of practical solutions. He adopted the 5W + 1H (why, what, who, where, when, and how) model, combined with literature review and expert knowledge acquisition, namely, focus groups and interviews. This leads to (i) clarifying the general interrelationships among smart energy cities, smart cities, and sustainable cities; (ii) a holistic, multidisciplinary, and comprehensive smart energy city conceptual framework that reveals its principles, goals, intervention areas, stakeholders, and time and space scales; and (iii) a set of smart energy practical solutions and technologies, divided into eight intervention areas [5]. Badii believes that a large number of activities have been carried out in defining smart city architecture to cope with this complexity, and a large number of different types of services and processes have been implemented. He introduced the work carried out in the context of the Sii-Mobility Smart City Project. His project defines a smart city architecture that solves a wide range of processes and data. To this end, by proving the use of semantic ontology and knowledge base in data aggregation in smart service production, the latest solutions of smart city architecture based on data aggregation and smart city API are compared. Baddi's solution proposes to aggregate and re-coordinate data (open and private, static and real-time) through the use of reasoning/intelligent algorithms to achieve complex service interactions through smart city APIs. His work was developed in the context of the Sii-Mobility National Smart City Project, which combines mobility and transportation with smart city services [6]. Daniel believed that technological advancements in multiple fields have promoted the development of smart city applications, which can improve the lifestyle of modern cities. When using a visual sensor in this situation, still images and video streams can be retrieved from the monitored area, which may provide valuable data for many applications. In fact, visual sensor networks may require highly dynamic conditions to reflect changes in parameters in smart cities. In this case, the characteristics of the visual sensor and the conditions of the monitored environment, as

well as the status of other concurrent monitoring systems, may affect the way the visual sensor collects, encodes, and transmits information. He proposed a fuzzy-based method to dynamically configure the operating mode of the vision sensor in sensing, encoding, and transmission modes, using different types of reference parameters. This innovative method can be considered as the basis for multisystem smart city applications based on visual surveillance, and is expected to bring significant results to this research field [7]. Smart cities ensure the sustainable economic and social development of cities by improving the efficiency of urban governance and operation. As each city faces different development foundations, the most important thing in the initial stage of smart city construction is to find a development model suitable for local characteristics. Therefore, all localities must face up to their own level of informatization development and choose a suitable road for smart city construction.

Research on the role of science and technology development in promoting cities: the development of cities mainly relies on the Internet, through which the Internet of Things and the Internet are closely combined to achieve technological innovation and application. Compared with the urban sustainability framework, the modern technology and "intelligence" in the smart city framework have received more attention. Based on the two major hidden dangers of the robo-advisory service platform's regulatory risk control dilemma and investor irrational emotions, this study innovatively proposes a platform governance analysis framework of "platform risk supervision + investor decision-making capabilities" [8]. From the two paths of platform supervision and investor irrational sentiment, build a robo-advisory service platform risk prediction model and an investor's emotional decision-making model for the robo-advisory service platform, so as to conduct targeted governance on the robo-advisor service platform from different perspectives. In addition, this research analyzes theories of financial development, technological innovation, and smart city construction, including the theoretical basis of financial development, technological innovation, and smart city construction [9–11], as well as the analysis of the mechanism of action among financial development, technological innovation, and smart city construction.

2. Fintech Platform Risk Assessment Regulation Algorithm Method

2.1. The Module Division and Overall Architecture of the Robo-Advisory System. Robo-advisor generally refers to robot financial management. Robot financial management is the introduction of artificial intelligence into traditional financial advisory services, instead of physical robots helping customers manage financial affairs. Instead, it provides automated investment portfolio recommendations through computer program algorithms based on the investment objectives and risk tolerance set by the demander through online interactions. Unlike traditional counter-counter face-to-face financial management services, which require many service personnel, the purpose is to improve efficiency. The

robo-advisor is divided into four major functional modules, as shown in Figure 1. The first functional module is the data collection module, which is responsible for the collection of Exchange Traded Fund (ETF) transaction data and the calculation of the recent rate of return [12]. The data flow from the information collection layer, and its main function is to collect analysis data related to the assessment object and the information technology risk, including assessment task information, basic information of banking business system, bank information technology management, business process status investigation, as well as vulnerability threats and asset value identification based on business system assets. ETF mainly involves three participating entities, namely promoters, trustees, and investors. The promoters and fund product founders are generally stock exchanges, large fund management companies, and securities companies. The trustee is entrusted by the promoter to custody and control all the assets of the stock trust portfolio. Since the index ETF adopts an indexed investment strategy, unless the index changes, the trustee generally does not need to adjust the stock portfolio from time to time, but the trustee of the management investment company ETF has a certain degree of freedom of investment decision-making power. Trustees are generally financial institutions such as banks, trust, and investment companies. The second functional module is the user management module, which is responsible for the management of user accounts, the management of user risk tolerance ratings, and the management of user investment portfolios. The third functional module is the portfolio recommendation module, which is responsible for the preprocessing of transaction data, the calculation of the recommended portfolio (completed through the portfolio model), and the storage of the recommended portfolio. The fourth functional module is the monitoring and rebalancing module, which is responsible for simulating the value trend of the investment portfolio, monitoring the deviation of the portfolio fund allocation ratio during the calculation of the trend, and adjusting the ratio back to the recommended state when the deviation exceeds the threshold.

The article studied the supervision of robo-advisors, but as far as the existing robo-advisors papers are concerned, more research focuses on the fields of finance and information engineering. Therefore, in the writing process, the research results of robo-advisors in economics, finance, and other related disciplines are integrated, starting from law or economic law, and understanding robo-advisors from as many angles and dimensions as possible.

The main logic processing and calculation work in robo-advisors are concentrated on the server, and the front-end is mainly responsible for sending requests and processing a small amount of data, which are a typical application type suitable for B/S architecture [13–15]. Based on this, the entire system is divided into three parts: data storage, server, and Web front-end. The overall system architecture is shown in Figure 2. Among them, the data storage part is responsible for storing user information, historical ETF transaction data, the recent daily rate of return of each ETF, and information about recommended investment portfolios and user investment portfolios. The server is responsible for receiving
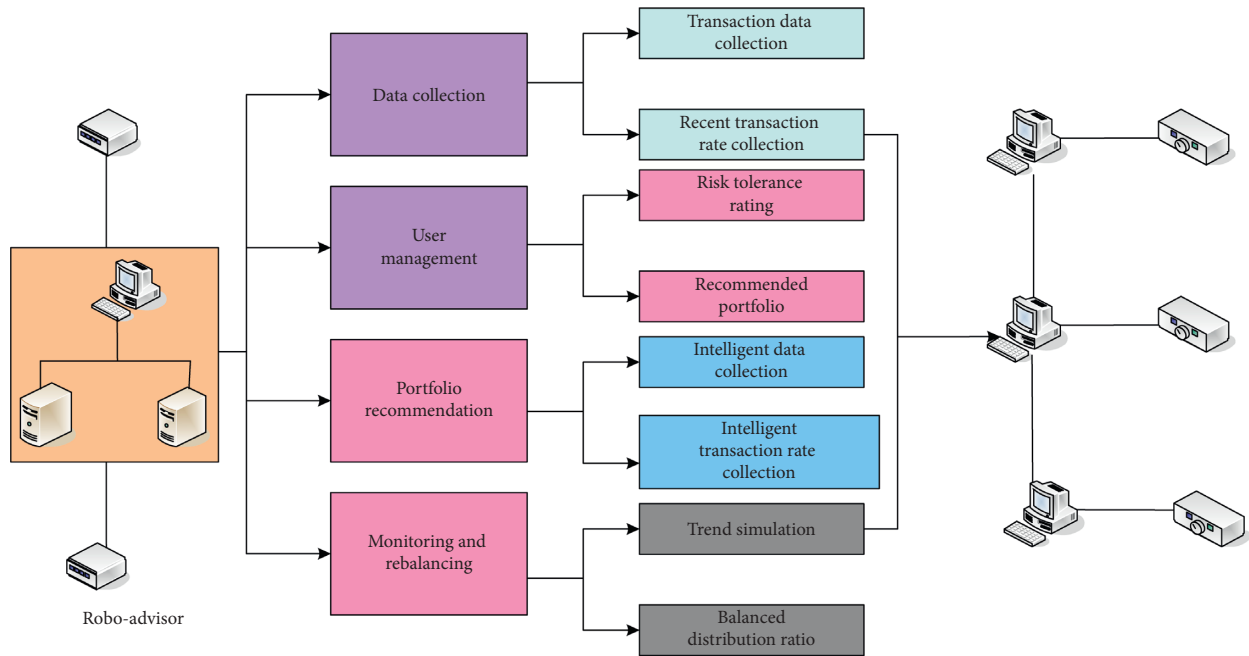
FIGURE 1: Module division of the four major functional modules.
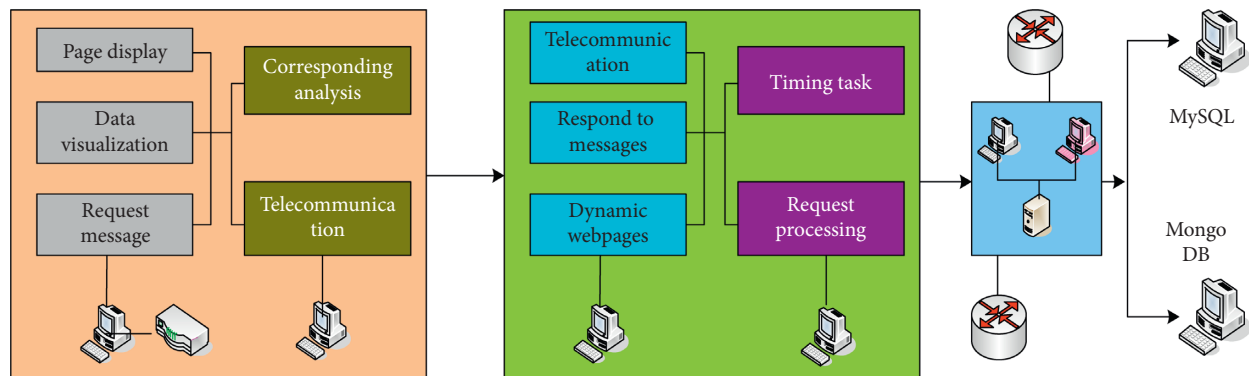
FIGURE 2: Overall system architecture.

front-end requests, performing logical processing, reading and writing data, and sending responses to the front-end. The web front-end is responsible for providing users with operating interfaces, sending requests to the server, and receiving and displaying the results returned by the server to users.

2.2. Adaptation Module Design. Intelligent analysis is divided into user analysis and investment analysis. User analysis is to analyze according to the evaluation questionnaire filled by the user during the user data acquisition stage, to classify different customers, and to allocate the proportion of large-scale assets according to the results of the rating.

When securities companies develop robo-advisor projects, they collect the "views" of retail customers through the analysis of retail customer historical data, market peer experience, and massive transaction data mining, complete the establishment of customer characteristic factors, and analyze and summarize the customer characteristic factor library. And through the subdivision of factors, it can be summarized into categories such as customer's overall characteristics, transaction characteristics, investment style, investment strategy, investment skills, and position status, including multiple characteristic factors. The process of collecting customer characteristic factors is shown in Figure 3.

At the same time of user analysis, investment analysis is also carried out. Investment analysis is based on the analysis of fund data obtained in the data acquisition stage. Investment analysis is divided into two steps. The first step is to select the investment target. When selecting the target, it can first determine the broad categories of assets, and select several outstanding funds of various assets to join the asset pool through data analysis. The specific choice can be based on fund company ranking, fund manager ranking, fund performance ranking, and some fund evaluation indicators, such as Sharpe ratio and maximum retracement of the
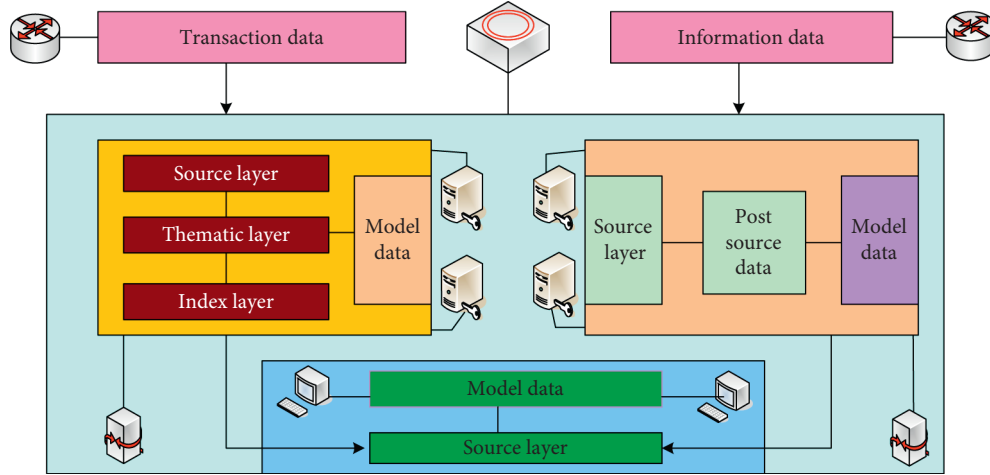
FIGURE 3: The process of collecting customer characteristic factors.

interval, through these methods to select the best funds. A comprehensive analysis of the customer's historical investment characteristics and investment skills will be carried out, and the customer's trading operations will be recorded in the customer's characteristic factor system, and the investment characteristic factors of the company's investment consultants will also be collected [16]. The characteristic factors of the two include industry selection, individual stock selection, profit and loss ratio, high-frequency trading period, timely stop loss rate, timely profit settlement ability, risk control ability, industry preference, position control, individual stock preference, and other characteristic vectors. Using the similar formula calculation method in big data, the vector distance between the customer and the investment adviser is calculated, and the distance represents the similarity between the two. The size of the distance and the level of similarity are positively correlated, that is, the closer the distance, the higher the similarity between the two. For comparison, the system will focus on recommending top five investment advisors with matching degrees to customers, showing their views, simulated investment portfolios, investment advisory products, and self-selected stocks. In strict accordance with the relevant measures for the management of industry investment advisors, investment advisors can provide clients with securities advisory services and investment advisory portfolio products.

The calculation method of matrix element $D(i, j)$ in the similar formula calculation method in big data is as follows:

$$D(i, j) = d(i, j) + \min\{D(i - 1, j), D(i, j - 1)\}, \quad (1)$$

where $d(i, j)$ represents the distance between the i-th point in the sequence X and the j-th point in the sequence Y.

For the similarity distance of two time series:

$$\Delta D = \min\{D(i - 1, j - 1), D(i, j - 1), D(i - 1, j)\},$$
$$D(i, j)_\beta = d(i, j) + \Delta D. \quad (2)$$

In addition, the stock market is very time-sensitive. Most of the research on the historical trend of stocks is to have a better prediction of the future. It is hoped that similar stocks will have

similar trends in the future. Although the future trend of stocks cannot be fully estimated, it is possible to make appropriate guesses from recent information dynamics. Therefore, the more similar the two stock sequences are in the later time, it means that the two time series are more similar. Therefore, this research improves the similarity calculation method, introduces exponential weighted moving average, enhances the similarity of recent patterns, and weakens the similarity weight far away from the current time. The weighted formula is as follows:

$$D(i, j) = \beta(i, j) + (1 - \chi)\min\{D(i - 1, j), D(i, j - 1)\}, \quad (3)$$

where β is the weight coefficient.

The formula for attractiveness is as follows:

$$X(m, n) = s(m, n) - \max\{a(i, k) + s(i, k)\}, \quad (4)$$

where (m, n) is the similarity between points m, n.

In the iterative update, similar clustering of big data may oscillate due to some noise data, so the smoothing factor α is introduced for control, which is also called the damping coefficient, and the previous trend is considered in the current trend direction. The formula for one iteration of attraction X and belonging A is as follows:

$$X = (1 - \alpha)X_{T-1} + (1 - \lambda)X_{T-1},$$
$$A = (1 - \alpha)A_{T-1} + (1 - \lambda)A_{T-1}. \quad (5)$$

The similarities between stocks are described in different ways, and then the stocks are divided into categories by splicing different clustering algorithms. Exploring the pros and cons of the similarity of different methods for stock time series data, and the differences between different clustering algorithms, select the most suitable clustering algorithm for stock classification, and enhance the accuracy of stock classification.

In Markowitz's modern portfolio theory, indicators to measure returns and risks are proposed:

$$R = \sum_{i=1}^{n} wr, \quad (6)$$

R is the expected return of the asset portfolio.

The formula for portfolio risk is as follows:

$$\lambda^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \lambda_{ij}. \tag{7}$$

Establishing an effective investment portfolio by finding the minimum portfolio standard deviation under the condition of a given expected return R:

$$\min \lambda^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \lambda_{ij} + \gamma,$$

$$\text{STR} = \sum_{i=1}^{n} w_i r_j. \tag{8}$$

Using the risk parity model to optimize the combination formula is as follows:

$$\min f = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[w(\Omega w)_i - w(\Omega w)_j \right]^2. \tag{9}$$

The formula for calculating the consistency CI is as follows:

$$\text{CI} = \frac{(\lambda - N)}{(N - 1)},$$

$$\text{RI} = \frac{(K - N)}{(N - 1)}. \tag{10}$$

where N is the order of the judgment matrix.

Calculation consistency CR:

$$\text{CR} = \frac{\text{CI}}{\text{RI}}. \tag{11}$$

2.3. Portfolio Recommendation Module.

Existing investment advisory platforms generally use two methods for user analysis to determine the investment ratio. One is to give a few fixed investment ratios for users to choose, such as the Betterment website. This result is actually more inclined to the user's subjective risk perception and ignores the user's objective conditions. It is very likely that the investment results will not match the customer's own situation. The second method is to divide the product portfolio's effective frontier in equal proportions to match the user's risk level. The result of this is not rigorous, and the difference in the choice of products in the effective frontier will also cause great deviations in the configuration results.

The portfolio recommendation module is the core of the entire robo-advisor. It is responsible for preprocessing the recent returns of each ETF, calculating the recommended portfolio through the portfolio model calculation, and storing the recommended results to disk. The optimized investment portfolio model is adopted in the implementation. The preprocessing of the recent returns of each ETF, including the processing of missing data items, and the calculation of the mean and covariance matrix of the returns based on the data. First, take the JSON string of the recent

return rate of each ETF from the MongoDB collection return-recent and convert it to the Data Frame type in the memory to obtain the return rate of 82 ETFs in the last 200 trading days. If there are missing items, it means that the corresponding single ETF stopped trading on the trading day due to some reasons of its own. Therefore, the return rate of the day is 0, and 0 can be directly added to the missing items. After dealing with the missing items in the data, it is necessary to calculate the mean return rate and the covariance matrix. Using Pandas to realize the EWM method to calculate the statistical characteristics of the mean and variance of the rate of return, and then use them as the input of the portfolio model, and run the model to get the recommended portfolio. The calculation of the portfolio model takes a lot of time, and it takes several hours in the serial case. The model needs to calculate the number of ETFs contained in the portfolio $K = 6, 7, 8, 9, 10$, respectively. Therefore, parallel operation can be realized for different K values. Because of the existence of the global interpreter lock in the Python language, only one thread can execute at any time, so the parallel implementation based on the multiprocess is adopted in the implementation, and the running time of the parallel implementation of the portfolio model has dropped to about half an hour.

There is a certain correlation between financial development and technological innovation:

$$\text{LNP} = C + \lambda_1 \text{SCID} + \lambda_2 \text{GS} + \lambda_3 \text{LNF}, \tag{12}$$

SCID is the smart city development index.

The dynamic panel model is as follows:

$$\text{SDI}_{\text{IT}} = \lambda_1 \text{SDI} + \lambda_2 \text{SDI}_2 + \lambda_3 \text{Lnp} + \lambda_4 \text{Finance} + \lambda_5 \text{Control},$$

$$\text{SDI} = \lambda_1 \text{LNP} + \lambda_2 \text{Control} + \gamma. \tag{13}$$

Among them, SDI and SDI_2 represent the SDI index lagging one and two periods, respectively.

2.4. Database Implementation.

The database of this evaluation system uses SQL-Server. In the implementation process, in order to facilitate the replacement of other database software in the future, this system uses SQL standard statements instead of calling SQL-Server specific functions. Designing a database is to classify tables according to dictionary tables and data tables. Among them, the dictionary table is used to define the data used in the coding. These data have been defined as rules at the beginning of the evaluation, such as asset importance assignment table, fragile assignment table, and asset type, and basically, no addition, deletion, and modification operations are performed. Data tables are often added and deleted during the evaluation process, such as asset identification tables, business system tables, and risk management tables. The asset value assignment information is shown in Table 1.

2.5. The Design of the Display Module of the Robo-Advisor System.

The robo-advisory system displays a "360-degree stereo view" of customers. The comprehensive investment ability of customers is scored through multiple dimensions

TABLE 1: Asset value assignment information.

Field name	Type of data	Length	Describe
ID	Int	4	Identification number
Name	Nvarchar	25	Asset materiality description
Rate	Int	4	Value of the asset
Description	Nvarchar	230	Description information

such as risk control ability, investment ability, ability to grasp opportunities, and operation ability, and the score is compared with the average score of stock customers to form a "customer profile," and fully mine customer transaction data to obtain customer investment habitual preference factors. Opening the page of the robo-advisor system to display the timing ability and risk control ability of the customer's investment ability analysis, and display the score and the average score of securities customers. It also displays customer characteristics obtained through testing: turnover rate, preferred industry, relevance of investment and research reports, preferred investment products, preferred risk, degree of concentration or divergence of holdings, and long, medium, or short-term investment risks, etc. Asset risk information is shown in Table 2.

After the investment advisory system configures the investment ratio, feedback is required. The result of the investment portfolio analyzed by the intelligent algorithm is fed back to the customer. The content includes the expected return rate, return volatility, investment asset types and their allocation ratios, and investment recommendations. The investment proposal states the basis of the investment portfolio and the customer's risk tolerance. Customers can invest according to investment recommendations.

Regarding how to recommend suitable financial products to suitable customers, it is assumed that the influence conditions of customers choosing to purchase financial products are customer investment experience, risk level, capital turnover rate, capital scale, expected rate of return, subscription threshold, investment cycle, and investment strategy. Through the simulation system, the purchase rate of different financial products, the order in which customers with the same capital scale purchase products and the product category with the highest purchase probability for customers with the same risk level are calculated. The intelligent model self-learns, trains, and tests the model to construct the most suitable model scheme. At the same time, the weight adjustment of subjective weighting method and the screening rules of customer suitability conditions are added to form a product recommendation list.

The similarity rate of rise and fall in the same category refers to the probability that stocks in the same category will rise and fall in the same direction for a period of time in the future. This article uses SDI to represent the same direction similarity rate of rise and fall, and its calculation method is as follows:

$$\text{SDI} = \frac{\sum_i^N \max\{I_1, I_2\}/m}{N}. \tag{14}$$

Among them, N is the number of categories.

The similarity rate of rise and fall in the same direction Z_D is as follows:

$$Z_D = \frac{\alpha * \text{SDI}/\lambda}{(\text{PVI} + \beta) * (\lambda + \beta)} = \frac{\alpha * \text{SDI}}{\lambda (\text{PVI} + \beta) * (\lambda + \beta)}. \tag{15}$$

Let Q denote the optimal value of the second-stage problem, which is related to the random parameters and the first-stage decision X.

$$\begin{aligned} Q &= \min q(y, \xi), \\ w(y, \xi) &= h(y, \xi) - G(y, \xi). \end{aligned} \tag{16}$$

Then the stochastic programming model S_G of the compensation problem has the following form:

$$S_G = \min f(x) + E(x, \xi). \tag{17}$$

Nested optimization model:

$$\begin{aligned} W(\xi)Z_1 &= H_1(\xi_1) - G(\xi_1), \\ W_T(\xi)Z_T &= H_T(\xi_1) - G(\xi_1)Z_{T-1}, \end{aligned} \tag{18}$$

where ξ is a random variable.

3. Results of Risk Assessment and Regulation of Financial Technology Platforms

Financial technology companies rely on advanced technology to settle in the market competition, and improper disclosure of internal technical personnel of the company, vulnerabilities in the internal technical system of the company, and improper release of corporate products and services that may all cause the leakage of key corporate technologies. Once the technology is leaked, imitated, or plagiarized by market competitors, what the company loses is not only the waste of the huge investment in research and development in the early stage but also the core competitiveness and potential business opportunities for survival in the market. This blow to enterprise development may be devastating. The main risks of financial enterprises are shown in Figure 4.

Due to various historical reasons, the financial industry will not be linked with "good user experience" and "focus on user needs" in people's vision. On the contrary, it will give people the impression of scarce resources and high service attitude. However, with the impact of the Internet, companies that pay more attention to convenient operation and timely service information will be more favored by users. With the development of Internet finance, according to statistics, it is estimated that in 2021, the number of online payment users will reach approximately 1.33 billion. The payment method statistics are shown in Figure 5. In recent years, the introduction of fast payment in more and more industries shows that the upper limit of the development of smart cities is more far-reaching than people think. Through the introduction of technical means, the convenience of urban residents' life has been further improved, and the use of scientific and technological means has given the government more choices in urban governance and prevention.

Table 2: Asset risk information.

Field name	Type of data	Length	Describe
ID	Int	4	Identification number, primary key
Relevance ID	Int	5	The vulnerability threat association table ID is set as a foreign key related to the vulnerability threat association table
Device ID	Int	6	Corresponding to the asset ID in the vulnerability threat association table t_ relevance
Threat ID	Real	4	Corresponding to the vulnerability ID in the vulnerability threat association table t_ relevance
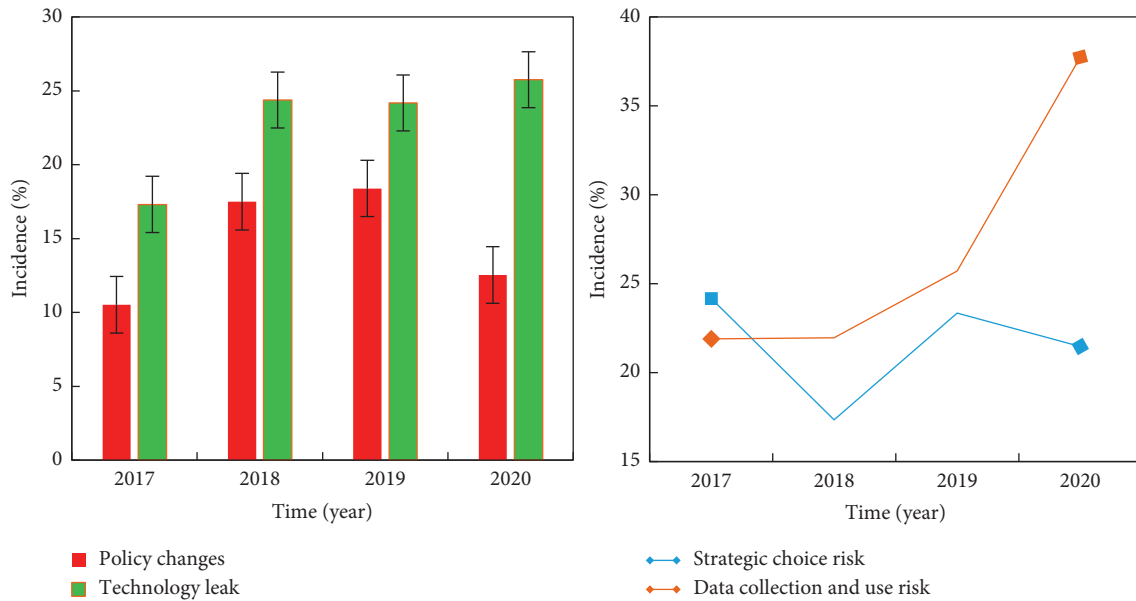
Figure 4: Main risks of financial companies.

For the spot silver from January 01, 2017 to June 05, 2017 on the precious metals exchange in City A and the spot silver 9995 on the precious metals trading center in City B, three K-line data of 60 min, 180 min,, and 240 min were used, use the similar formula calculation method in calculating big data to test the conventional accuracy of parameters, and simultaneously monitor the resource utilization of the server. Part of the result data is shown in Table 3.

During the test, the tested function of the system did not run abnormally under the system constraints, which affirmed the guiding role of the similar formula calculation model in the big data on the spot real market. Judging from the test results, the algorithm in this study has successfully made a profit for 25376 times in a total of 30564 transactions, and the accuracy rate is as high as 83.04%.

Based on the above, the risk values of the three major risk domains of bank information technology risk are 54.05, 51.63, and 64.28, respectively. The three risk domains are weighted by collecting opinions from multiple parties, and the final weighted risk assessment results are shown in Table 4.

When clustering data, it is necessary to calculate the distance between different time series separately. Usually, when calculating the distance, the base of the series is different, which will cause large errors. These errors sometimes lead to bizarre data results. The similar formula calculation

method in big data is used to solve such problems and speed up the calculation. The timing trend before and after standardization is shown in Figure 6.

The higher risk values in the management domain are information security (19%), information system development, testing and maintenance (21%), information technology operation (30%), and business continuity management (30%). The comparison of risk values in the management domain is shown in Figure 7.

Under the management support center, there are departments such as affairs, finance, project management, and the head of the department reports to the head of the management support center. There is no separate compliance department, intellectual property department, risk control department, and internal audit department that directly report to the board of directors. All compliance, intellectual property, risk control, and internal audit functions are currently concentrated in the legal department. The company's main intellectual property rights, personnel status, and financial status in the past three years are shown in Table 5.

Experiments found that a single LSTM and AVG algorithm is difficult to reduce the loss of the model, and under-fitting occurs. Therefore, the LSTM and AVG algorithms cannot effectively explain the investor emotional decision model. Due to the enhanced feature extraction
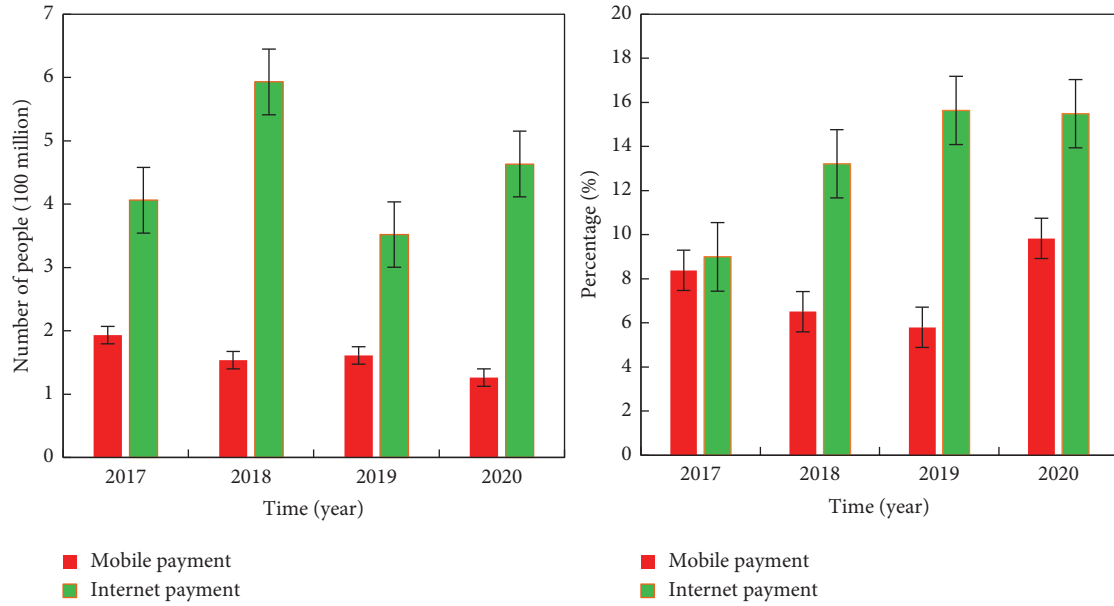
FIGURE 5: Payment method statistics.

TABLE 3: Part of the result data.

Variety name	Number of successful long signals	The number of successful short signals	Short-selling success rate
1	131	131	0.8771
2	111	13	0.8193
3	101	10	0.8193
4	101	13	0.8713
5	103	107	0.81
6	110	13	0.8783
7	117	11	0.8187
8	107	13	0.8731

TABLE 4: Final weighted risk assessment results.

Risk domain name	Risk domain weight	Risk domain risk value	Risk domain weighted risk value
Management domain	0.3	54.05	10.88
Business process domain	0.3	51.63	15.78
Business system domain	0.5	64.38	33.64

ability of the risk assessment regulation algorithm and the algorithm of this paper, which integrates Attention, it has a better fit to the investor's emotional decision data, the loss continues to decline, and achieves a relatively stable result. Among them, the risk assessment regulation algorithm model in this study makes the final loss smaller. The loss reduction degree of the four algorithms is shown in Figure 8.

The accuracy of various algorithms is shown in Table 6, which shows that the accuracy of the two-way LSTM and the similar formula calculation method in the fusion of big data has been greatly improved on the basis of a single LSTM and AVG network, by 17% and 10%, respectively. Although the similar formula calculation method in big data integrated with Attention has stronger feature extraction capabilities (with an accuracy of 88%), due to the lack of data on the robo-advisory service platform itself, there may be an overfitting problem. Therefore, the experimental results

show that the similar formula calculation method in big data can better explain the emotional decision-making ability of the robo-advisor service platform investors, which provide algorithmic support for the subsequent emotional calculation of platform investors.

Based on the similar formula calculation model in the fusion of big data with experimental training, this study uses Python language and Pytorch framework to calculate sentiment scores for the eight major platform reviews, takes the average value, evaluates the ranking, and compares it with the platform's own comprehensive strength ranking. The specific ranking and sentiment score are shown in Figure 9. The experiment found that investors' risk analysis and platform selection capabilities are still insufficient, and investors' emotional decision-making on the robo-advisor service platform has a large gap with the official platform's comprehensive strength. For example, Huitou Bank ranks
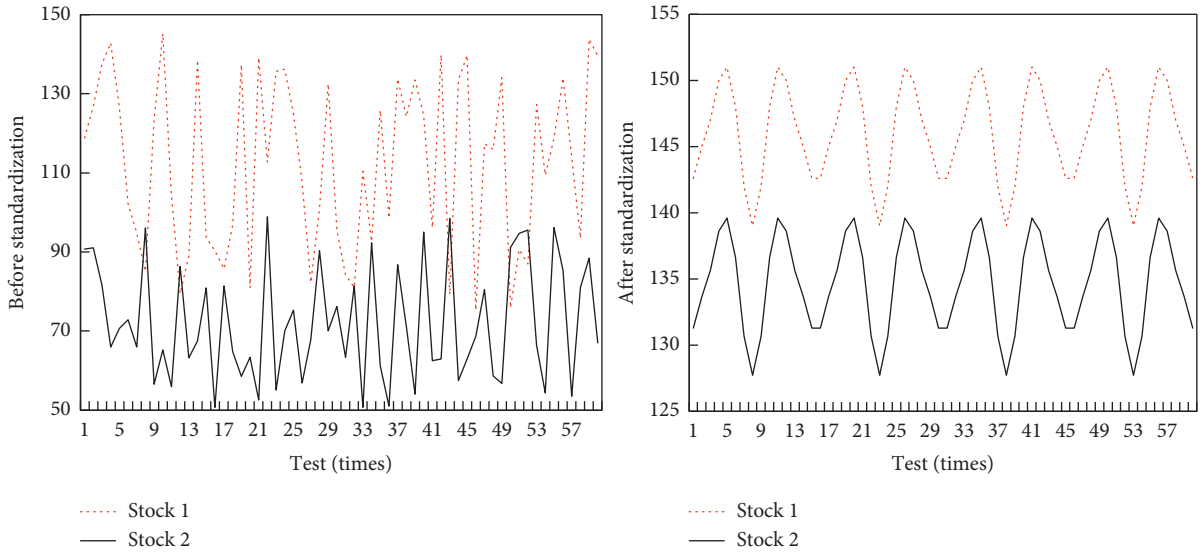
Figure 6: Time series trends before and after standardization.
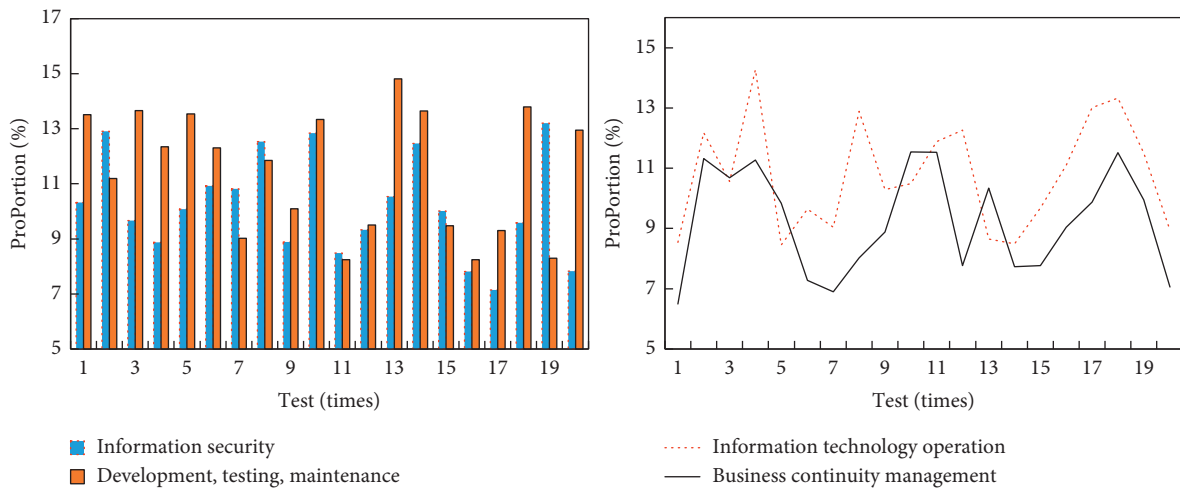
Figure 7: Comparison of risk values in the management domain.

Table 5: The company's main intellectual property rights, personnel status, and financial status in the past three years.

Year	Net assets	Sales revenue	Total profit
2019	24172.17	440.21	−2441.41
2018	26611.91	121.47	−2161.49
2017	29777.29	5.67	−10022.72

very high in the official comprehensive strength announcement, but investors are not optimistic about its evaluation, only reaching a score of about 0.2. It is comprehensively found that investors still lack the ability of rational risk identification and platform selection for robo-advisory service platforms. This experiment explores the characteristics of relevant investor comments on the robo-advisory service platform and the concerns and emotional characteristics of investors on the robo-advisory service platform. First, construct a domain dictionary, segment words, remove stop words, and extract keyword frequencies

from the cleaned comment data. It can be found that keywords such as "income," "risk," "loss," "handling fee," and "fund" appear more frequently, which are the main points that investors pay attention to. The statistical results of investors' attention to keywords are shown in Figure 9.

4. Discussion

In the era of "AI for all people," computers, as the basic tools for processing work, can help people deal with various tedious and repetitive tasks, and they also have a place in the
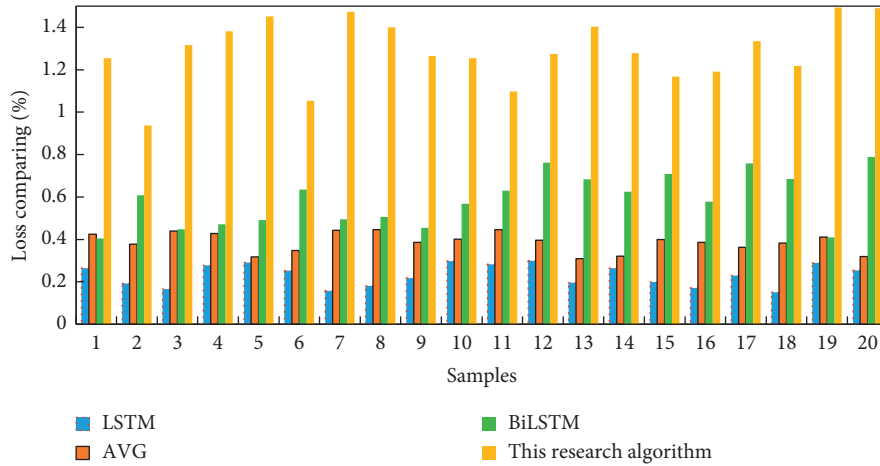
Figure 8: Loss reduction degree of four algorithm learning.

Table 6: Accuracy of various algorithms.

Model/evaluation index	Net assets (%)
AVG	70
Bidirectional LSTM	71
Similar formula calculation method in big data	86
Similar formula calculation method in big data fused with attention	88
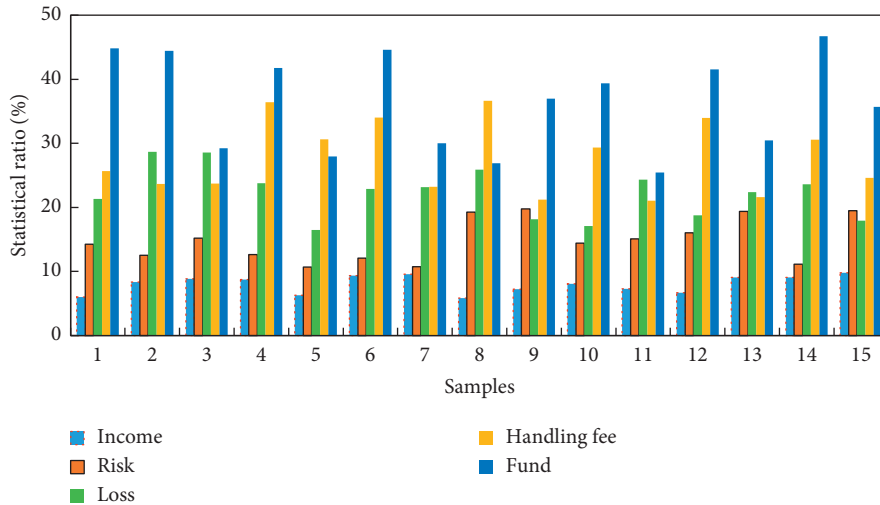
Figure 9: Statistics of investors' attention to keywords.

field of financial information analysis. By categorizing this information, it is convenient for people to obtain information that is useful to them. The main research content of this subject is to use computer combined with machine learning methods to classify stocks, and on this basis, to dig out the relationship between stocks or categories, so as to better help users to invest. The market environment of perfect competition often stays at the ideal level, and the realistic market economy often shows the characteristics of imperfect competition. Traditional trade theory believes that in the case of monopolistic competition, open trade can not only expand the scale of the market, deepen industry

competition, and increase the degree of international specialization of production, but it can also alleviate the distortion of resource allocation caused by monopoly, make the product price and cost lower, benefit consumers, and improve the national income level.

The financial market plays an important role in the development of the country and the national economy. There are many types of financial products in the financial field. Stock is the most important part of securities, and it is one of the most common investment choices in people's daily life. However, stocks have the characteristics of high risk and high return. Therefore, how to avoid risks and

obtain returns when investing in stocks has been a topic of more concern in the society in recent years. Therefore, when studying the relationship between stocks, the relationship between stock categories is considered. Investing in stocks of different categories can be more conducive to reducing investment risks, and the relationship between categories is often more beneficial to investment, because the relationship between many stock categories is often more stable and lasting than the relationship between individual stocks, and a stable relationship can better avoid risks in investment.

Platform speculation and investor irrational sentiments are two important causes of platform risks and turbulence in the financial market environment. On the one hand, the robo-advisory service platform must provide investors with efficient and convenient robo-advisory services, and on the other hand, it must also restrain the uncertain speculative behavior in it and prevent the occurrence of tragic phenomena similar to a large number of P2P online loan investment platforms. This also brings challenges to the government and regulatory authorities. When the regulators provide incentives for robo-advisor concepts to enter China, they are also constantly exploring how to prevent and control the risks of bad investment platforms. For example, the current operation of the robo-advisory service platform needs to obtain the qualification certificate of the China Securities Regulatory Commission, but there are still the phenomena of platform speculation, running off or losing assets of investors.

For example, blockchain technology is based on the concept of "decentralized distributed transparent mechanism," which replaces the one-way linear relationship of "intermediary–customer" and introduces a highly democratic principle of decentralization into the financial market. Under the above background, the limitations of a single regulatory body, a single regulatory means, and a single regulatory model have become more and more obvious. In the context of financial technology, financial regulatory agencies will fall into a long-term trilemma that they can only achieve two of the three goals of seeking clear rules, maintaining market credibility, and encouraging financial innovation. "The drive of the information regulation model to the power of social regulation" provides a new way of thinking for the socialization and diversification of financial technology regulation. Therefore, information regulation can complement administrative regulation to achieve functional advantages, integrate regulatory resources from the government, market and society, and build a multilevel compound model of financial technology regulation.

The change from Internet finance to financial technology means that regulatory activities are becoming more complex and professional. However, with the rapid development of technical means, technology is subtly changing the way and value of information regulation, and it is an inevitable trend to reshape information regulation. In the context of financial technology, the content and structure of information regulation have become more and more complex, and instrumental rationality and value rationality have shown a dualistic and potentially unified contradiction in information regulation. As the current Internet financial market is undergoing a period of fierce changes and strong supervision, and a large number of Internet financial platforms are closed or cleared, the market is in a special period of transition from "Internet finance" to "financial technology." Investors still lack the ability of robo-advisory service platform selection and risk identification, and there are two key points in its governance: on the one hand, investors' irrational decisions account for a large proportion. Platform operators and the government or the Securities Regulatory Commission can enter from this to cultivate investors' sense of rational investment, including the popularization of relevant financial investment knowledge and attention to important risk characteristics. On the other hand, it is more important to supervise the compliance operation of the platform and restrict its disclosure of relevant characteristic information. For example, investors are concerned about the income, profit and loss, product structure, etc., and review them, which promotes the sound development of robo-advisors in China and builds a benign investment environment for investors.

Monitoring the problems and deficiencies in the financial markets of various regions, by introducing interactive items, helps to examine the degree of financial development's effect on technological innovation, and can discover and evaluate the degree of integration of financial development and technological innovation in various regions, check whether local financial institutions are in place in implementing the state's policy on financial support for innovation and development. On the basis of having a clear understanding of the financial development and technological innovation of various regions, it helps all regions to proceed from the actual situation and in-depth exploration of the specific path and the best way of the integration of technology and finance. Practicing the spirit of the "18th National Congress" promotes the realization of the goal of an innovative country by enhancing the role of finance in technological innovation. With the continuous in-depth research on smart cities, the focus of research is gradually starting from the nature and connotation of urban construction from what is a smart city and how to build a smart city. It not only evaluates the sustainability of urban construction but also pays more attention to the sustainable development of people. With the diversification and standardization of research, smart cities in particular have now developed into a national strategic plan; how to use better methods to construct and how to evaluate the performance of a complex system of smart city financial projects has become a research object that needs to be focused on nowadays.

5. Conclusion

Driven by digitization and informatization, financial technology plays an important role in promoting financial decentralization and financial democratization. The use of scientific and technological means and Internet technology to provide convenient and accurate inclusive financial services for long-tail customers has realized the decentralization and multicentralization of power in the financial

market. This article's research on Internet financial information regulation is still mainly based on the context of "Internet finance." The research on the paradigm of information regulation in the context of "financial technology" is still insufficient, and the research on the characteristics of the integration of information regulation and technology needs to be in-depth. In addition, the study of relevant Internet financial judicial cases still needs to be further excavated and sorted out. This study combines the latest developments of robo-advisory products, comprehensively analyzes the actual situation of robo-advisory platforms, and constructs a model of influencing factors about the use behavior of robo-advisors. This study analyzes the investor's emotional decision-making ability based on the information asymmetry theory and the bounded rationality theory of economic behavior. The existing domestic regulatory policies may not be able to effectively supervise robo-advisory service platforms. In such an environment, robo-advisory services are also difficult to meet the needs of investors. One of the important reasons for this is the irrational investment sentiment of investors. At present, investors' investment philosophy is not yet mature; for example, investors generally value short-term returns.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Acknowledgments

References

[1] J.-Y. Yeh and C.-H. Chen, "A machine learning approach to predict the success of crowdfunding fintech project," *Journal of Enterprise Information Management*, 2020.

[2] H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluagac, A. Kadri, and A. Tuncer, "UAV-enabled intelligent transportation systems for the smart city: applications and challenges," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 22–28, 2017.

[3] T. Bo, C. Zhen, G. Hefferman et al., "Incorporating intelligence in fog computing for big data analysis in smart cities," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2140–2150, 2017.

[4] F. G. Brundu, E. Patti, A. Osello et al., "IoT software infrastructure for energy management and simulation in smart cities," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 832–840, 2017.

[5] F. Mosannenzadeh, A. Bisello, R. Vaccaro, V. D'Alonzo, G. W. Hunter, and D. Vettorato, "Smart energy city development: a story told by urban planners," *Cities*, vol. 64, no. APR, pp. 54–65, 2017.

[6] C. Badii, P. Bellini, D. Cenni, A. Difino, P. Nesi, and M. Paolucci, "Analysis and assessment of a knowledge based smart city architecture providing service APIs," *Future Generation Computer Systems*, vol. 75, no. oct, pp. 14–29, 2017.

[7] C. Daniel, C. Mario, P. Giovanni, and D. Cristian, "A fuzzy-based approach for sensing, coding and transmission configuration of visual sensors in smart city applications," *Sensors*, vol. 17, no. 1, pp. 1–19, 2017.

[8] Y.-H. Yuan, S.-H. Tsao, J.-T. Chyou, and S.-B. Tsai, "An empirical study on effects of electronic word-of-mouth and Internet risk avoidance on purchase intention: from the perspective of big data," *Soft Computing*, vol. 24, no. 8, pp. 5713–5728, 2020.

[9] H. Yue, H. Liao, Li Dong, and L. Chen, "Enterprise financial risk management using information fusion technology and big data mining," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 3835652, 13 pages, 2021.

[10] Z. Xu, G. Zhu, N. Metawa, and Q. Zhou, "Machine learning based customer meta-combination brand equity analysis for marketing behavior evaluation," *Information Processing & Management*, vol. 59, no. 1, Article ID 102800, 2022.

[11] C. Gu, "Application of data mining technology in financial intervention based on data fusion information entropy," *Journal of Sensors*, vol. 2022, Article ID 2192186, 10 pages, 2022.

[12] Y. Chen, W. Zheng, W. Li, and Y. Huang, "Large group Activity security risk assessment and risk early warning based on random forest algorithm," *Pattern Recognition Letters*, vol. 144, pp. 1–5, 2021.

[13] O. I. Khalaf and G. M. Abdulsahib, "Optimized dynamic storage of data (ODSD) in IoT based on blockchain for wireless sensor networks," *Peer-to-Peer Networking and Applications*, vol. 14, no. 5, pp. 2858–2873, 2021.

[14] S. Namasudra and P. Roy, "PpBAC," *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 14–31, 2018.

[15] Y. Zhang, H. Huang, L.-X. Yang, Y. Xiang, and M. Li, "Serious challenges and potential solutions for the industrial Internet of things with edge intelligence," *IEEE Network*, vol. 33, no. 5, pp. 41–45, 2019.

[16] A. M. Al-Momani, M. A. Mahmoud, and M. S. Ahmad, "Factors that influence the acceptance of Internet of things services by customers of telecommunication companies in Jordan," *Journal of Organizational and End User Computing*, vol. 30, no. 4, pp. 51–63, 2018.

Research Article

Distributed Simulation System for Athletes' Mental Health in the Internet of Things Environment

Baoyan Fu ⓘ **and XinXin Fu**

Jiangxi Teachers College, Yingtan, Jiangxi 335000, China

Correspondence should be addressed to Baoyan Fu; 2003026016@jxsfgz.com

Psychological troubles in training competitions mainly include worry about mistakes, long-term lack of improvement in sports performance, and lack of confidence in competitions. The main troubles in daily study and life are future career development and life planning, injury and illness, insomnia, and poor emotional control. Athletes are interested in psychological skills training, hobby training, interpersonal communication and other coaching content to improve sports performance. Athletes tend to prefer one-to-one psychological counseling and group counseling activities; there are differences in the psychological distress, coping styles and expected psychological counseling content of athletes in different age groups and events. This paper firstly introduces the important role of psychological quality education in modern competitive sports. The influencing factors of athletes' psychological quality were analyzed. At the same time, combined with relevant practical experience, it starts from various perspectives and aspects such as improving the scientific literacy of coaches and building a harmonious atmosphere for training and competition. This paper puts forward some effective strategies to strengthen athletes' psychological quality education and improve sports performance. In addition, it expounds the author's understanding of this, hoping to contribute to the practice of athletes' psychological quality education.

1. Introduction

In modern competitive sports, the intensity and difficulty of competition are getting bigger and bigger, which puts higher demands on the physical and mental health of athletes. Under the current situation, we must pay attention to the important role of psychological quality education in the application of modern sports. Objectively examine the many influencing factors affecting the mental health of competitive athletes, grasp the key points of the mental quality education of athletes, and promote the optimization and improvement of sports performance [1]. Based on the current situation of the needs and deficiencies of health monitoring and management of various types of athletes, athletes need to maintain a good mental state during training and competition. This is because a good mental state is conducive to its maximum potential and bear greater pressure. Generally speaking, they can be divided into three categories. The first

category, such as smart bracelets or watches, has the function of automatic health data collection, but lacks follow-up analysis and connection with health services. The second category, such as online health analysis and diagnosis platforms, can generate health reports based on big data, but data collection is not automatically connected to no health services. The third category, such as fitness or nutrition recommendation apps, passively provides health services, such as fitness courses, but is not based on intelligent recommendation and drainage based on data analysis [2]. Therefore, there is an urgent need for a digital platform that integrates real-time collection and monitoring of personal health data, automatic generation of health analysis reports, and follow-up intelligent recommendation to connect with health services. It is based on the above-mentioned status quo, to carry out research on the application of the Internet of Things and artificial intelligence to realize a health monitoring and management platform that spans time and
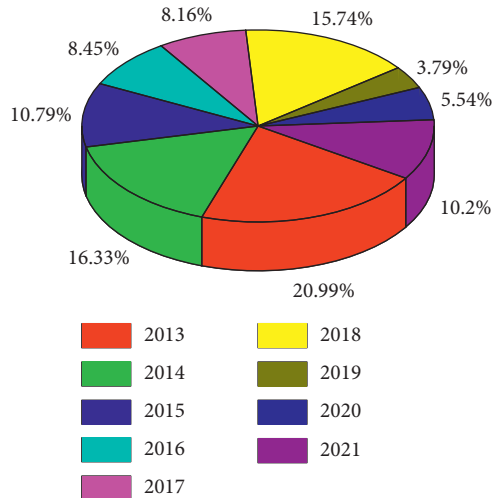
Figure 1: The occurrence of mental problems in athletes over the years.

space [3]. Based on the above analysis of the status quo and deficiencies, there is currently no platform or product that connects health data detection and subsequent health recovery services. Therefore, the *IoT* cutting-edge technology that integrates the Internet of Things, artificial intelligence deep learning and intelligent recommendation algorithms is proposed to realize the three-in-one process of "data based health monitoring and management service platform research based on the Internet of Things and deep learning" [4]. Research on health monitoring management service platform is significant for maintenance of athletes' health. When athletes use the platform, they can achieve accurate collection of health data and real-time update and view of health data through *IoT* collection devices. In addition, it will also recommend a series of health service management such as one-to-one live teaching by fitness coaches, one-to-one consultation with professional nutritionists, tweeting services on health forums, and Internet + diagnosis and treatment services based on the athletes' current physical health data. The health monitoring and management service platform based on the Internet of Things and deep learning can serve everyone who cares about their own health, and is precisely positioned for athletes with fierce competition and difficult psychological construction. Automatic real-time health data detection of users through IoT devices, big data analysis, and intelligent generation of user health monitoring analysis reports [5]. As shown in Figure 1, they are also students with rapid physical and mental development, and they will also face the problems that ordinary teenagers will encounter in their growth, such as academic pressure and internal contradictions in physical and mental development. The two influence each other and have their own particularities. In the previous International Olympic Committee, experts also emphasized that while paying attention to the psychological quality of athletes' competitive performance, it is necessary to pay attention to the mental health problems that often occur in athletes [6–9].

The full use of information technologies such as the Internet of Things in athletes' service management can effectively improve the technical quality and safety assurance of athletes' mental health management services. This approach can effectively enhance the high importance of effective management of their own health for all types of athletes. While effectively reducing the cost of psychological medical expenses, it can also change the bad mental life and behavior of athletes, and can effectively prevent the occurrence of mental diseases in athletes. In this way, it can effectively promote the recovery of athletes after various diseases, and further improve the economic, material and spiritual life of college teachers and students [10]. The Internet of Things project includes infrared sensors, laser scanners, and information sensors. According to the relevant protocol requirements, the connection between items and the Internet is realized, and the information flow between them is realized, so as to achieve the purpose of monitoring, tracking, identification, positioning and management. With the gradual increase in the number of sports and competitions in our country, there are higher requirements for venues, security and athlete training facilities. With the development of the Internet of Things project, it has brought more interactivity to the development of sports in our country, and has a very good development situation. It can be said that the development of the Internet of Things project. It will definitely bring significant changes to the development of sports in our country [11]. Athletes' mental health issues are getting more and more attention. Previous researchers have generally used tools such as clinical screening tools or mood state scales from a 90-item symptom list to assess the mental health of athletes. Although these tools are widely used and have shown good reliability and validity in many groups, their use in the field of sports has exposed the difficulty of daily monitoring and screening of sports teams and the irrelevance of the content of the scale to the sports context. Rice developed the Athlete's Psychological Stress Questionnaire from the perspective of the most obvious psychological stress on the development of athletes' mental health symptoms for daily monitoring and screening of professional athletes' mental health status [12]. Psychological stress refers to a state of feeling stressed and difficult to cope with, and is a state of constant emotional exhaustion and reaction to stressful experiences. Compared to internalized symptoms that other clinical diagnostic tools focus on, psychological stress can screen athletes for early signs of possible mental health problems for early professional support [13, 14]. In recent years, there have been more and more mental health problems among adolescents, which is worrying and has attracted attention from all walks of life. The relevant functional departments of the state, provinces and cities have also issued documents on the mental health of young athletes for many times, and the research on the mental health of young athletes is also in-depth. However, there are relatively few studies on the mental health of young athletes. Based on this, this study starts from the Internet of Things technology to investigate the mental health of athletes in youth sports training centers, establishes a health analysis system model. The real-time stream of data from IoT sensors, combined with historical data from other
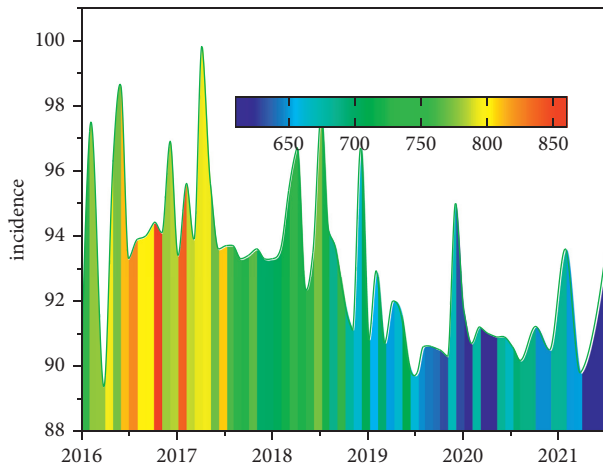
FIGURE 2: Percentage of athletes with mental problems over the years.

FIGURE 3: Psychological evaluation score curve of participating athletes.

projects, can be used to monitor the current work site and provide an ever-increasing set of data that can be used for simulating system to monitor athletes' mental health. This study aims to investigate the mental health status of athletes based on the Internet of things technology, establish a health analysis system model, and compare the effects of different factors on Athletes' mental health by horizontal and vertical simulation.

2. A Brief Introduction to the Internet of Things

The concept of the Internet of Things was first proposed by the *MIT Auto-ID* Research Center in 1999. This technology is a technology that can use radio frequency identification technology to identify and manage the researched affairs and laws anytime and anywhere under network conditions, which is what we call the interconnection between items. The essence of the technology is mainly to effectively integrate radio frequency identification (*RFID*) technology and the Internet, and to effectively apply the integrated products. *RFID* is a wireless communication technology that can use radio signals to identify a specific target and read and write relevant data without the need to establish mechanical or optical contact between the identification system and a specific target.

As shown in Figure 2, the emergence of the Internet of Things can effectively solve the interconnection between objects to objects, people to people, and people to objects. The Internet of Things is also quite different from the traditional Internet. For example, the interconnection between people and objects we are talking about mainly refers to the connection between people and objects through the use of some common devices. The interconnection between people mainly means that the process of interconnection between people no longer only depends on the connection method of personal computers [15].

$$E_k = 0.5 \sum_{i=1}^{q} \left(Y_i^k - C_t^k \right)^2. \tag{1}$$
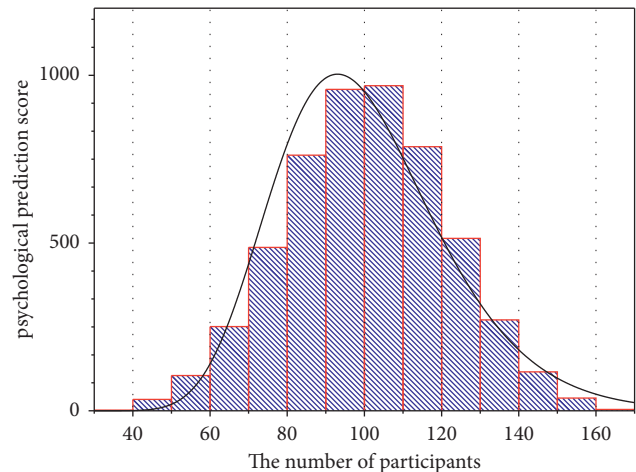
In the formula, E_k represents the error between the expected value and the actual value of the neural network.

The application of the Internet is based on optical fiber, to form a network of multiple computers to realize the communication and sharing of information. However, the network environment established by the Internet is actually virtual, and people can only apply the virtual information existing in it, but cannot change the form of real objects in the world. The emergence of *IoT* technology is to remedy this problem. At the same time, the Internet of Things can also realize the unmanned driving of vehicles, so that they can automatically avoid vehicles and congested roads. Thereby avoiding the occurrence of other situations such as traffic jams and saving more time for people. Based on this, we can find that the application of Internet of things engineering has high requirements for relevant equipment and devices. It needs a certain effectiveness and stability to improve people's work efficiency and quality of life [16].

The Internet of Things realizes the Internet between objects, through sensing devices, *RFID* technology, and identification technology. As shown in Figure 3, according to relevant protocols, the network connection between objects is realized, and the information exchange between objects is realized, so as to achieve the purpose of intelligent tracking, positioning, management and monitoring. This intelligent system will be combined with media, services and enterprises in the future to form the future Internet model [17]. The exact methods of data processing based on the SPSS software are data trend analysis, data contrast analysis and data subdivision analysis.

The essence of the Internet of Things is a convergent application in the development of information technology, and its characteristics mainly include the following aspects. First of all, the Internet of Things has the characteristics of the Internet, and it needs to be connected to the Internet before the intercommunication between objects can be realized. Secondly, a supervision cloud platform for environmental protection functional departments is established. Finally, the Internet of Things must be intelligent. The
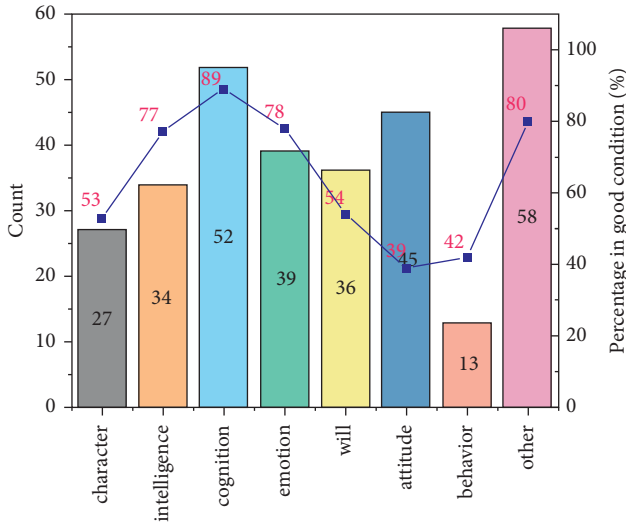
FIGURE 4: Psychological and behavioral impact levels.

FIGURE 5: Health status of male and female athletes over the years.

intelligence mentioned here refers to self-feedback ability, intelligent control ability and automatic application ability. According to the above description, the application of the Internet of Things project in the mental health of athletes will greatly promote the development of sports in my country. In sports, it is necessary to bring out the potential of people's physical, psychological, athletic ability and physique, so it is necessary to formulate a series of training programs and methods [18].

$$F_i = \frac{(x_i - \overline{x})}{\overline{x}} \times 100\%. \qquad (2)$$

In the formula, F_i represents the deviation of the index from the mean value of the index group. The larger the value, the higher the mean value, and the lower the mean value on the contrary.

At present, most of the health management work in my country is carried out by physical examination institutions and medical institutions, but in actual work, their focus is disease control rather than health management in the true sense. As shown in Figure 4, with the popularization of the concept of the Internet of Things, some people are now focusing on monitoring various indicators of the human body in real time through the use of computer and mobile phone software. As a result, the factors threatening the mental health of athletes are simulated and analyzed in the early stage, and effective solutions are proposed. Make a huge contribution to the development of the sports industry. The Internet of Things has become a game changer for all sports. The IoT sensors are now being used to measure minute details of all aspects of performance to help athletes improve their performance, such as running, swimming, cycling and so on [19, 20].

3. Operation of Mental Health Management Based on Internet of Things

According to the market survey and relevant statistical results of relevant research institutions in our country, it is
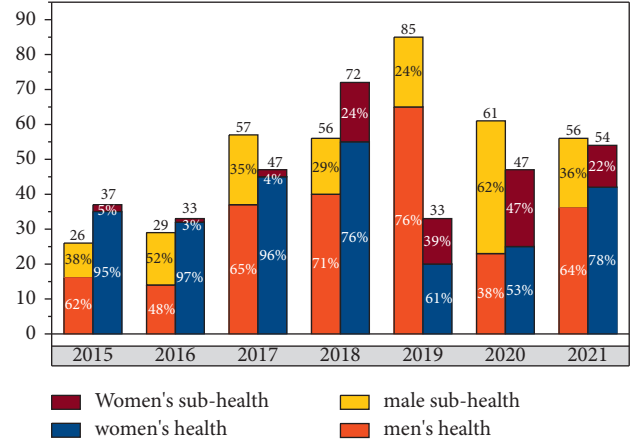
known that in recent years, the mental and physical quality of various athletes in many countries around the world needs to be further improved. This is because with the rapid progress of the people, the competitiveness in sports competitions is also increasing. Correspondingly, it will also bring huge psychological pressure to many athletes, so many relevant institutions and organizations have made great efforts to increase the importance and degree of education on athletes' mental health literacy. Mental health test management needs to include a lot of test content, among which athletes' mental health management test and physical fitness evaluation management are very critical parts. Therefore, to seriously and effectively improve athletes' physical and psychological quality, we must work hard. Create a rational model for the management of mental health testing of athletes [21, 22]. The continuous real-time stream of data from IoT sensors, combined with historical data from other projects, can not only be used to monitor the current work site. However, it also provides a set of data that can be used for machine learning for predictive analysis to make athletes' mental health monitoring system smarter.

3.1. Correct Choice of Frequency.
According to the performance classification and characteristics of its operating frequency, electronic tags can be divided into low frequency (*LF*: 30–300 kHz), high frequency (*HF*: 3–30 MHz), ultra-high frequency (*UHF*: 300MHz-3 GHz) and microwave frequency band (*MW*: 2.45–5.8 GHz) four frequency bands. *RFID* applications in the international are generally based on *LF* and *HF* tags as the main implementation. As shown in Figure 5, at this stage, the relevant institutions has mastered the core design technology of *HF* chips, and has already put it into actual mass production applications on a large scale. The data in Figure 5 comes from public data resources published on the Internet. Meanwhile, the *UHF* chip has also successfully completed the research and development process and development stage [23].

3.2. Privacy and Security Protection Measures.
RFID product labels have their own unique characteristics, but in fact they

should all have a unified product code based on the global electronic product label code standard or a unique product code in the world. And this product label has been written into a corresponding product label before their packaging work is completed or before they leave the factory, and it is not allowed to read and write settings at will. At the same time, all the databases in the core chip are fully encrypted in parallel and have passed international security certification indicators, which can fully ensure the security of data sources and processed information. This can not only effectively prevent the artificial cracking of encrypted links and other databases, but also automatically lock some of our important confidential data information according to the sector mechanism, further ensuring that our data security will not be leaked.

$$\Delta v_j = -\alpha \left(\frac{\partial E_k}{\partial v_j} \right) = \alpha d_i^k b_j. \tag{3}$$

In the formula, the correction value of the network weight is Δv_j, which represents the influence strength of the neural network error change on the output result.

3.3. System Cost Analysis.

In the context of the social Internet, the cost of electronic tags has become lower and lower. Since 2006, the price of each tag has dropped to 5 cents. At the same time, because the tag can realize the functions of reading and writing data, and can also be recycled many times, the cost of the tag becomes lower and lower.

$$\Delta w_{ij} = -\beta \left(\frac{\partial E_k}{\partial w_{ij}} \right) = \beta E_j \alpha_i. \tag{4}$$

In the formula, $\triangle w_{ij}$ represents the change value of the connection weight from the hidden layer to the input layer.

3.4. International Standard Issues.

At present, the 2009 edition of *RFID* and other related information technology and electronic equipment development and application standard specifications and related data information of the standard are being researched and formulated internationally. These sources mainly use the International Organization for Standardization (*ISO/IEC*), and currently the main standards for my country are all based on the 2009 edition of *ISO/IEC* 15693. Basically, the specific research or drafting of the industry technical standards related to the national standards of the industry has been successfully completed. With reference to the relevant national standards formulated by *ISO/IEC*1800 in 2009, their standards have also been officially included in the industry standard national key project implementation plan.

$$\Delta \theta_j = -\beta \left(\frac{\partial E_k}{\partial \theta_{ij}} \right) = \beta E_j. \tag{5}$$

In the formula, θ represents the threshold change of the hidden layer.
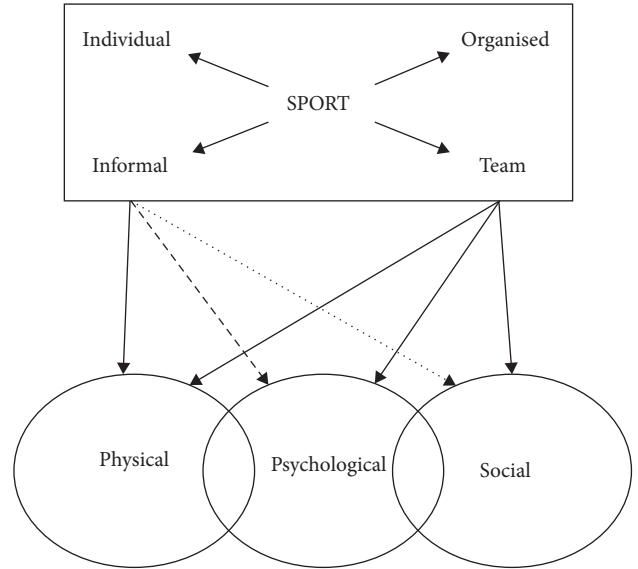
FIGURE 6: The relationship between mental and social health in sports.

3.5. Implementation of System Design Plan.

Take advantage of some current hospital management systems and professional local area networks. Such as the *HIS* and *LIS* of large universities, institutions and government public administration departments, and on this basis initially complete the treatment of hospital patients, equipment and medical equipment. The system is mainly used in clinical medicine, constitutes the main part of the system, and is the system hardware and related software for optimizing the management of commonly used drugs. We can see from Figure 6 that the most important parts of the hardware include *RFID* network tags, *RFID* network antennas, file readers, *RFID* operating systems, servers, terminals, and related network cables. And each component is mostly connected in larger shapes by different network devices, all of which are equivalent to a broad network running an operating system. However, software functionality generally refers to the software applied to all *RFID* tag servers and tag users for each mobile station.

Through the use of the Internet of Things method, the real-time management of sports equipment can be realized, thereby helping athletes to reasonably allocate time and methods during exercise. We can see from Figure 7 that real-time monitoring of the warehousing and delivery of sports equipment, record the application of equipment and equipment, and share the records to improve the application efficiency of sports equipment. In this way, it can effectively improve the training efficiency of athletes and greatly reduce their psychological burden.

$$p(y|x) = \frac{e^{yx^T w}}{1 + e^{yx^T w}}. \tag{6}$$

This formula represents the mathematical definition of logical distribution under this model, which is a method to complete the classification based on the discriminant of logical distribution.
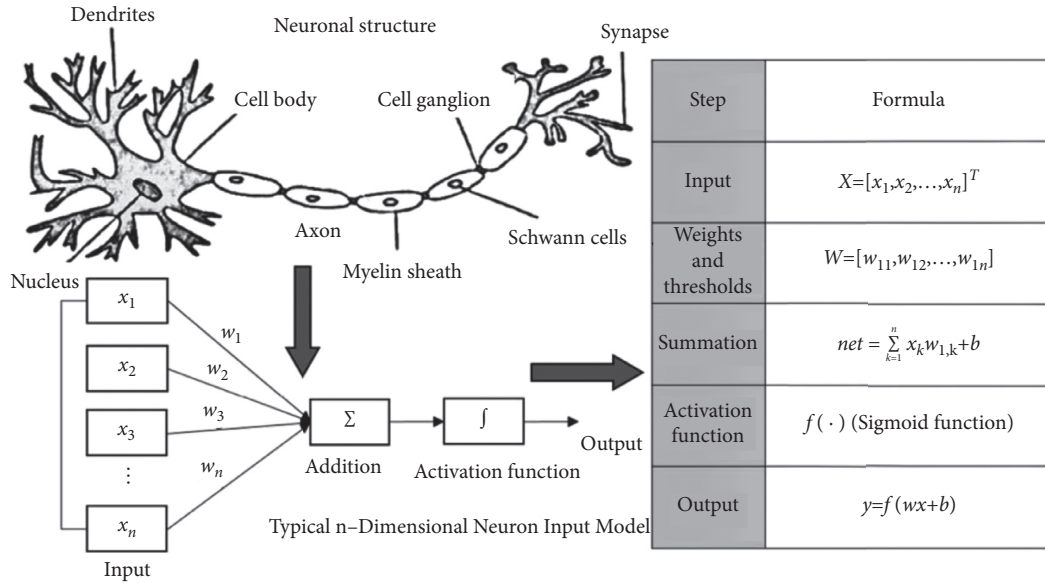
Figure 7: System structure diagram.

4. Application of Internet of Things Engineering in Sports

Using the Internet of Things project to monitor sports is actually monitoring the physical conditions of athletes in real time, because these physical conditions can reflect psychological changes to a certain extent. The terminal data collection method is used for information analysis. Once abnormality is found in the relevant parameters, an alarm will be issued immediately, thereby reminding the coaches and team doctors to take corresponding psychological counseling to avoid the occurrence of sports accidents during training. We can see from Figure 8 that this monitoring method can be applied to sports training as well as competitive competitions. Athletes can use the Internet of Things to monitor their physiological parameters during the usual training process. For example, for sprinters, they need to have strong explosive power, and in the process of sprinting, the heart will instantly enter. In the state of high-speed operation, how to train to improve the ability of the heart is very important.

$$f(w; X, y) = L(w; X, y) + R(w). \tag{7}$$

The above formula represents a model that rewrites the objective function in the form of a loss function.

Compared with training, athletes will continue to be in a state of high-intensity exercise in competitive competitions, and we can sometimes see athletes die suddenly during competitions. So far, hundreds of athletes have died suddenly during the competition due to too intense exercise. A very important reason for this phenomenon is that athletes do not pay attention to psychological adjustment during high-intensity training. If real-time monitoring of organs in athletes can be carried out, athletes can be warned before organ problems occur. Thereby, it can stop exercising, and effectively improve the psychological counseling, thereby
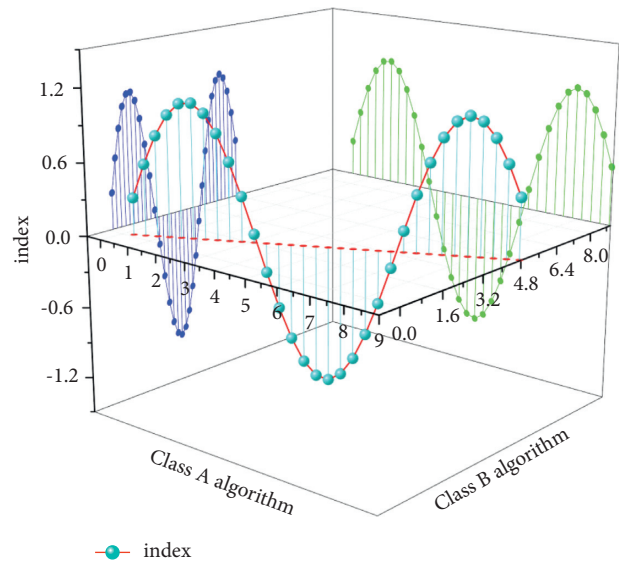
Figure 8: Comparison of the two algorithms.

avoiding the phenomenon of excessive exercise of athletes. Athletes cannot directly perceive their own mental state, so IoT engineering is needed to provide technical support.

$$c_j^* = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i^j. \tag{8}$$

In the formula, the number of mental health data objects in the c_j cluster is n_j.

As the continuous development of Internet cloud computing technology and various hardware facilities, some equipment systems using the Internet of Things cloud technology have emerged. This is a health management system that includes various subsystems such as various biological detection instrument systems, wireless signal

transmission and reception systems, terminal data processing systems, and background database storage and processing systems. In S mode, the data on the server can be added, modified and deleted through their own terminal machine, and the analysis results and exercise intervention plan fed back to the terminal by the background server can also be obtained. Users can exercise according to the plan. The data during the exercise will be transmitted to the background server to further guide psychological counseling.

$$J^* = \sum_{k=1}^{n_j} \sum_{j=1}^{k} \left\| x_k^j - c_j^* \right\|^2. \tag{9}$$

If $|J^* - J| < \xi$, it means that the clustering criterion function converges, and the final clustering result is obtained, and ξ is the criterion for judging the convergence of the function.

The mental health monitoring work process of athletes generally generates a huge amount of data, and the analysis of these health monitoring data has extremely important practical significance. More work mistakes make the mental health monitoring work of the athletes lose its value.

$$A_i = \frac{\sum_{i=1}^{3} a_y \partial_y}{\sum_{i=1}^{3} y}. \tag{10}$$

In the formula, A_j is the value corresponding to each feature layer index calculated, and the coefficient of each base layer is determined, and the value is 0–1.

The popular science service provided by the athlete mental health tracking system is no longer the original one-way indoctrination, but it is changing to two-way interaction, that is, the two-way communication between the disseminator and the recipient of the popular science service is realized by using the product of the athlete mental health tracking system as the communication medium. In addition, while the athlete mental health tracking system provides popular science and services to all sports circles, these health data can become the basis for the investigation of athletes' mental health quality. The athlete mental health tracking system is a service-oriented popular science system, which provides a good opportunity for the development of future popular science exhibits. The design of this system fully reflects the people-oriented concept, and the design is more in line with the actual needs of athletes advantages of the Internet.

5. The Important Role of Psychological Quality Education in Sports Competition

With the development of economy and society, sports competition has been continuously improved and has become a key part of human social life. In the complex, ever-changing and exciting competitive arena, only with stable and strong psychological quality can we better play our due level and achieve better sports results. In sports, the contest between athletes is not only a contest of their own technical level, but also a contest of psychological quality. When
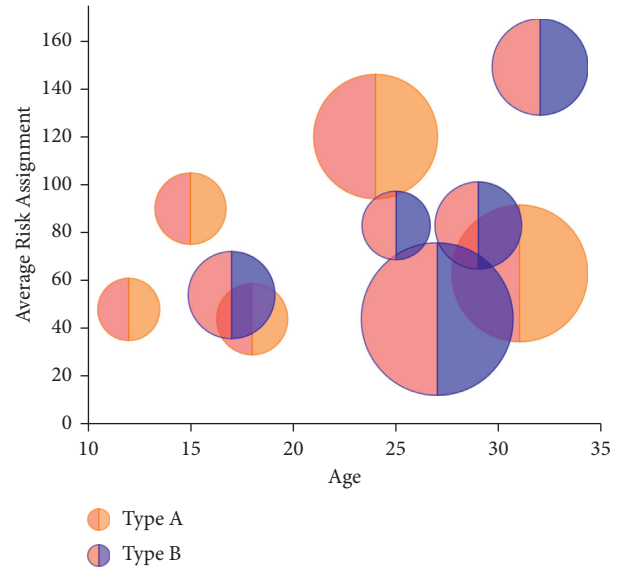
FIGURE 9: Average risk assignment for different ages.

athletes of the same level compete on the same field, the ups and downs of psychological state become the decisive factor for the performance of athletes.

Mental health is a key component of a person's overall quality. The level of modern competitive technology has generally improved, and the audience's viewing ability has also undergone profound changes. We can see from Figure 9 that if the mental health quality is lacking, it is bound to be difficult for the athletes to better adapt to the competition rhythm. Practice shows that the importance of athletes' psychological quality in their comprehensive quality is becoming more and more prominent, and it plays a pivotal role in competitive performance. Therefore, improving the level of psychological quality education and strengthening the psychological quality of athletes can better face failure and success, joy and tears.

Competitive sports under the commodity economy and social environment presents many new features, of which economy and utility are typical representatives. Under the impact of economic interests, some athletes pay too much attention to their own performance, put too much energy into utilitarianism, and even distort their self-values and outlook on life. Development requirements, unable to realize self-worth. Therefore, in the economic society, it is necessary to make athletes face fame and fortune more correctly and adapt to the needs of social development through systematic psychological education.

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}. \tag{11}$$

Among them, f_o represents the observed frequency; f_e represents the expected frequency.

Today's sports competition is no longer a purely physical competition, but a comprehensive competition in psychology, training, economy, technology, anti-stress and other aspects. The traditional one simply relies on infinitely extending training time and improving training intensity
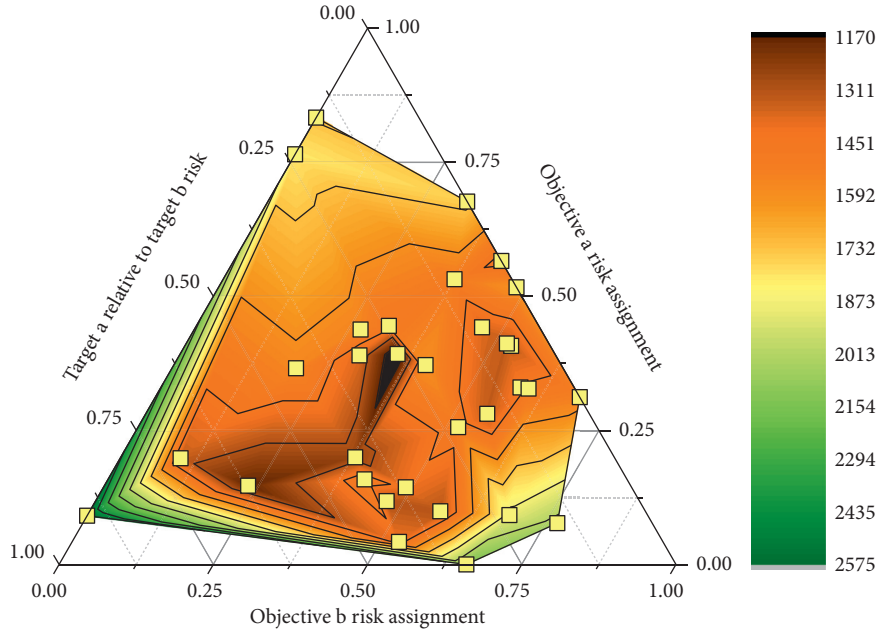
Figure 10: Risk value comparison of two targets.

The method is out of date. At the same time, the age requirements for training and participating in competitive athletes tend to decrease, and their psychological development is relatively lagging behind. All of the above conditions put forward higher requirements for the overall quality of the coaches.

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log \frac{p(x, y)}{p(x)p(y)}. \quad (12)$$

Among them, $p(x, y)$ represents the joint distribution of two random variables (X, Y) of the observed frequency; $p(x)$ represents the marginal distribution; $p(y)$ represents the marginal distribution.

6. Application of BP Neural Network in Intelligent Analysis

First, the wavelet transform is used to denoise the athletes' psychological monitoring data, and the denoised data is used as the input of the *BP* neural network. The number of nodes in the input layer depends on the dimension of the data source degree, such as the athlete's mental fitness index plus gender, age, height, and weight. The number of output layer nodes is 3, which are poor, medium, and excellent. The number of hidden layer nodes affects the learning time and training effect. We can see from Figure 10 that the formula is used to determine the number of nodes. Finally, make sure the number of nodes is 6. The network performance is relatively stable. At the same time, the threshold and weight in the network structure are used as the objective function, and the error value between the expected output and the predicted output is used as the fitness function. The optimal solution is obtained by continuous optimization calculation.

In this paper, wavelet analysis is used to denoise, to keep the wavelet coefficients of the effective signal to the greatest extent, and remove the wavelet coefficients of noise. At the same time, in the process of denoising, the smoothness and similarity of the original signal should be kept as far as possible, indicating that the variance estimation of the signal before and after denoising is the minimum value in the worst case. Wavelet analysis has strong denoising ability. The steps of the wavelet analysis method are to select the appropriate wavelet basis function, to decompose the original sample by wavelet, and to obtain the wavelet coefficients of each layer. The wavelet coefficients are processed by the soft threshold method to minimize the maximum mean square error of the estimated signal. The Internet of Things in sports shows the potential to collect and analyze data to improve the performance of spaces, things and athletes. Driven by the competitive nature of sport, iot has the potential to fine-tune its ability to take sports mental health monitoring to a new level.

7. Conclusion

For competitive sports, the ultimate goal of sports training is to create excellent sports performance, and the psychological quality of athletes is the most basic factor to improve competitive ability. After training with the training set, the *BP* network and the *PSO-BP* network are tested respectively using the data in the literature, which can accurately reflect the efficiency and accuracy of the *BP* neural network after particle swarm optimization. Aiming at the popularization of smart devices and the needs of athletes' training monitoring, a deep learning-based athlete's physique assessment algorithm was designed. The advantage of IoT technology is that the athletes can remove faults in time and reduce the occurrence of safety accidents in competitive sports. The

algorithm scores different athletes by selecting factors such as vertical jumping, high-speed leg raising, sitting and stretching, height, chest circumference, body fat percentage, and uses particle swarm optimization algorithm to optimize the *BP* algorithm to establish athlete psychology. The evaluation model of the situation comprehensively reflects the physical and mental quality of athletes. Through simulation experiments, it is concluded that the optimized *BP* network has high efficiency and accuracy. It is obviously superior to the traditional *BP* network in the evaluation and prediction of athletes' psychological quality.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Grassi, A. Vascellari, A. Vascellari et al., "Return to sport after ACL reconstruction: a survey between the Italian society of knee, arthroscopy, sport, cartilage and orthopaedic technologies (SIGASCOT) members," *European Journal of Orthopaedic Surgery and Traumatology*, vol. 26, no. 5, pp. 509–516, 2016.

[2] S. D. Barber-Westin and F. R. Noyes, "Factors used to determine return to unrestricted sports activities after anterior cruciate ligament reconstruction," *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 27, no. 12, pp. 1697–1705, 2011.

[3] D. T. van Yperen, M. Reijman, E. M. van Es, S. M. A. Bierma-Zeinstra, and D. E. Meuffels, "Twenty-Year follow-up study comparing operative versus n treatment of anterior cruciate ligament ruptures in high-level athletes," *The American Journal of Sports Medicine*, vol. 46, no. 5, pp. 1129–1136, 2018.

[4] B. B. Rosa, A. M. Asperto, C. P. Helito, M. K. Demange, T. L. Fernandes, and A. J. Hernandez, "Epidemiology of sports injuries on collegiate athletes at a single center [J]," *Acta Ortopédica Brasileira*, vol. 22, no. 6, pp. 321–324, 2014.

[5] F. Allhoff and A. Henschke, "The internet of things: foundational ethical issues," *Internet of Things*, vol. 1-2, pp. 55–66, 2018.

[6] E. C. Makhni, E. K. Crump, M. E. Steinhaus et al., "Quality and variability of online available physical therapy protocols from academic orthopaedic surgery programs for anterior cruciate ligament reconstruction," *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 32, no. 8, pp. 1612–1621, 2016.

[7] G. Arslan, "Loneliness, college belongingness, subjective vitality, and psychological adjustment during coronavirus pandemic: development of the College Belongingness Questionnaire," *Journal of Positive School Psychology*, vol. 5, no. 1, pp. 17–31, 2021.

[8] M. Buckthorpe, A. Frizziero, and G. S. Roi, "Update on functional recovery process for the injured athlete: return to sport continuum redefined," *British Journal of Sports Medicine*, vol. 53, no. 5, pp. 265–267, 2019.

[9] R. K. Britt, W. B. Collins, K. Wilson, G. Linnemeier, and A. M. Englebert, "eHealth literacy and health behaviors affecting modern college students: a pilot study of issues identified by the American college health association," *Journal of Medical Internet Research*, vol. 19, no. 12, p. e392, 2017.

[10] L. Lei, Y. Tan, K. Zheng, S. Liu, K. Zhang, and X. Shen, "Deep reinforcement learning for autonomous internet of things: model, applications and challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1722–1760, 2020.

[11] M. Harrer, S. H. Adam, R. J. Fleischmann et al., "Effectiveness of an internet-and app-based intervention for college students with elevated stress: randomized controlled trial [J]," *Journal of Medical Internet Research*, vol. 20, no. 4, p. e136, 2018.

[12] J. J. P. C. Rodrigues, D. B. De Rezende Segundo, H. A. Junqueira et al., "Enabling technologies for the internet of health things," *Ieee Access*, vol. 6, pp. 13129–13141, 2018.

[13] E. G. Lattie, E. C. Adkins, N. Winquist, C. Stiles-Shields, Q. E. Wafford, and A. K. Graham, "Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review," *Journal of Medical Internet Research*, vol. 21, no. 7, p. e12869, 2019.

[14] J. Kim, R. Choue, and H. Lim, "Differences of sebehavior, diet quality and quality of life in south Korean women according to their weight status," *Clinical nutrition research*, vol. 5, no. 3, pp. 161–171, 2016.

[15] J. T. Jost, "The marketplace of ideology: "Elective affinities" in political psychology and their implications for consumer behavior," *Journal of Consumer Psychology*, vol. 27, no. 4, pp. 502–520, 2017.

[16] D. Y. Lee, S. A. Karim, and H. C. Chang, "Return to sports after anterior cruciate ligament reconstruction - a review of patients with minimum 5-year follow-up," *Annals Academy of Medicine Singapore*, vol. 37, pp. 273–281, 2008.

[17] S. Grimmelikhuijsen, S. Jilke, A. L. Olsen, and L. Tummers, "Behavioral public administration: combining insights from public administration and psychology," *Public Administration Review*, vol. 77, no. 1, pp. 45–56, 2017.

[18] D. Zilio, "On the autonomy of psychology from neuroscience: a case study of skinner's radical behaviorism and behavior analysis," *Review of General Psychology*, vol. 20, no. 2, pp. 155–170, 2016.

[19] W. Wood and D. Rünger, "Psychology of h," *Annual Review of Psychology*, vol. 67, no. 1, pp. 289–314, 2016.

[20] P. Sheeran, J. A. Bosch, G. Crombez et al., "Implicit processes in health psychology: d," *Health Psychology*, vol. 35, no. 8, pp. 761–766, 2016.

[21] K. E. Childs, D. Kincaid, H. P. George, and N. A. Gage, "The relationship between school-wide implementation of positive behavior intervention and supports and student discipline outcomes," *Journal of Positive Behavior Interventions*, vol. 18, no. 2, pp. 89–99, 2016.

[22] B. Litwiller, L. A. Snyder, W. D. Taylor, and L. M. Steele, "The relationship between sleep and work: a meta-analysis," *Journal of Applied Psychology*, vol. 102, no. 4, pp. 682–699, 2017.

[23] J. S. Howard, M. L. Lembach, A. V. Metzler, and D. L. Johnson, "Rates and determinants of return to play after anterior cruciate ligament reconstruction in national collegiate athletic association division I soccer athletes," *The American Journal of Sports Medicine*, vol. 44, no. 2, pp. 433–439, 2016.

Hindawi

Retraction

Retracted: A Corpus-Driven Study on Three Elements of Third Personal Pronoun Endophora in English Abstracts of Chinese and Foreign Theses

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Luo, "A Corpus-Driven Study on Three Elements of Third Personal Pronoun Endophora in English Abstracts of Chinese and Foreign Theses," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5981285, 7 pages, 2022.

Hindawi

Research Article

A Corpus-Driven Study on Three Elements of Third Personal Pronoun Endophora in English Abstracts of Chinese and Foreign Theses

Yuxiao Luo (ORCID)

Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528402, Guangdong, China

Correspondence should be addressed to Yuxiao Luo; luoyuxiao@zsc.edu.cn

The corpus is applied to compare and analyze three elements of the third personal pronoun endophora in English abstracts of Chinese and foreign theses. It is discovered that there is no significant difference in the total frequency of third personal pronoun endophora used by Chinese and foreign masters, but there is a significant difference in three elements of third personal pronoun endophora. Native Chinese-speaking masters tend to choose "it" as the anaphor, while native English-speaking masters are inclined to select "them" as the anaphor; antecedents used by native Chinese-speaking masters are mostly consistent with keywords of theses, while antecedents used by native English-speaking masters are rarely consistent with keywords of theses; distances between antecedents and anaphors set by native Chinese-speaking masters are much longer than those set by native English-speaking masters. These findings are beneficial for Chinese masters in terms of better understanding the writing conventions of English abstracts of foreign theses and further improving the language quality of English abstracts in their theses.

1. Introduction

Endophora refers to the phenomenon that writers make use of words or phrases to substitute previous or subsequent elements. Endophora seems to be a simple language phenomenon, but actually it not only shows the substitution relation in grammar but also displays the coreferring relation in semantics. Moreover, endophora, as a cohesive means, plays the main constraint on cohesion. Halliday pointed out that anaphors are mostly enacted by personal pronouns [1]. Haiyan LU comparatively analyzed the usage of personal pronouns in general English writing by native Chinese-speaking students and native English-speaking students, and discovered that native English-speaking students tend to choose third personal pronouns as anaphors, while native Chinese-speaking students are inclined to select first personal pronouns to replace complicated and redundant antecedents [2].

In recent years, some Chinese and foreign studies have focused on comparing personal pronoun endophora in English sentences or general English discourses written by native Chinese-speaking students and native English-speaking students. According to Helen and Renate, it is learned that all types of personal pronouns are used as anaphors in general English writing, while in academic English writing, third personal pronouns rather than first and second personal pronouns are preferred [3]. However, the study on third personal pronoun endophora in academic English discourses is relatively rare. Compared with first or second personal pronouns, third personal pronouns not only follow the economic principle of language as they avoid repetitions of complicated antecedents, but also play the role of cohesive means as their comprehension relies on the discourse context. Therefore, third personal pronoun endophora is both a textual language phenomenon and a textual cohesive tool. It is currently a research topic in the spotlight for comparing and analyzing third personal pronoun endophora in academic English writing written by native Chinese-speaking students and native English-speaking students.

This paper employs the tool of corpus to study third-personal pronoun endophora in English abstracts of Chinese and foreign theses. On the basis of the analysis of similarities and differences in third personal pronoun endophora, it is expected to help Chinese masters better understand the writing conventions of English abstracts of foreign theses and to further improve the language quality of English abstracts in their theses.

2. Theoretical Framework

According to Halliday, endophora is complicated in that it needs to refer to some elements that have appeared or will appear in the discourse so as to make the whole discourse coherent. When the reference of an element is used to refer to its previous element in the discourse, this phenomenon is called anaphora; when the reference of an element is used to point to its following element in the discourse, this phenomenon is named cataphora [1]. In his dissertation, Jiangli XU claimed that anaphors and antecedents are two inherent elements of anaphora and cataphora [4]. Michael Strumpf demonstrated that anaphors in discourse are mostly expressed by objective third personal pronouns instead of subjective first and second personal pronouns [5]. Forms of personal pronouns consist of subject case form, objective case form, possessive determiner form, possessive pronoun form, and reflexive pronoun form. Compared with other forms, the possessive pronoun form owns double antecedents. For example, "hers" may have the antecedent of "Mary" and the antecedent of "book". Hua WANG described that those forms of third personal pronoun anaphors are as follows: "He," "His," "Him," "Himself," "She," "Her," "Herself," "It," "Its," "Itself," "They," "Their," "Them," and "Themselves" [6].

Antecedents in the discourse are generally served by nouns or noun phrases. Since the combination of nouns or noun phrases is infinite, the identification of antecedents is a complicated procedure. Chomsky attempted to propose a series of explanatory rules at the sentential level about antecedents' identification with the aid of the c-command theory of transformational grammar [7]. On the basis of the node analysis and tree-diagram presentation, Levinson claimed that antecedents that have a coindexed relation with anaphors can be identified in the discourse smoothly [8]. Fukumura also adopted c-command as the core idea and claimed that third personal pronouns usually have a coreferring relation with the definite noun phrase in independent sentences [9]. For instance, "Mary thinks she has been late." If this sentence is an isolated sentence, the third personal pronoun "she" will definitely be regarded as a variable bound by the definite noun phrase "Mary."

It is clear that Chomsky and Fukumura are restricted to the syntactic level. H.G. Widdowson strode the syntactic node to control the jurisdiction level and confirmed anaphora in discourses by analyzing the semantic characteristics of anaphors and antecedents [10]. However, his method requires a lot of time and energy to analyze the semantic features of third personal pronouns and noun phrases. In order to pursue the labor-saving principle, Levinson and Yan HUANG have raised the antecedent confirmation to the pragmatic level by analyzing partial interpretations, disjoint hypotheses, and salient information interfaces of the coreferring relation [8, 11], but two scholars never touch upon readers' cognitive motivation in the process of discourse comprehension. Ariel and Yulong XU classified anaphors into accessibility levels from the cognitive perspective, and they maintained that pronouns belong to high accessibility markers in that pronouns are closely related to the context in which short-term memory usually works [12, 13].

In comparison to anaphora, the reason why cataphora has not been extensively explored by scholars lies in its rareness. Halliday holds that cataphora is a special linguistic phenomenon which is often discovered in rhetorical or literary writing [1]. For instance, "Through their fiction, the Brontes consider Victorians' treatment and understanding of animals, as well as the human implications and consequences of cruelty." In this instance, the antecedent "the Brontes" appears after the anaphor "their", so this phenomenon is genuine backward anaphora. In addition to western scholars, Chinese scholars also make great efforts to investigate cataphora. Through analyzing materials, Yuan GAO and Canlong Wang discovered that although there indeed exists a third personal pronoun cataphora in news titles or subtitles, its application in English writing is still rare [14, 15]. For example, "It may not be great, but Footstep is not that bad". The anaphor "it" and the antecedent "Footstep" are coindexed in this news title. Therefore, in this research, much attention will be paid to the analysis and discussion of third personal pronoun anaphora instead of cataphora.

Those mentioned researches related to the third personal pronoun endophora indicate that endophora confirmation procedures in discourses should integrate all levels to form a comprehensive endophora confirmation principle. Specific procedures of endophora identification proposed by the author in this paper are as follows:

(1) The first step is to seek out the current topic before the anaphor. In the process of topic searching, pronouns are considered as high accessibility markers, so their antecedents are current topics.

(2) The second step is to judge whether the anaphor and the current topic are identical in semantic features or not. If they are identical, they have a coreferring relationship; if they are not identical, they do not have a coreferring relationship.

(3) The third step is to infer previous and subsequent antecedents based upon contexts at a pragmatic level, and specific contexts are of great importance to identify antecedents.

3. Research Design

3.1. Research Questions. Three research questions will be answered:

(1) What are the similarities and differences in the total frequency of third personal pronoun endophora in the English abstracts of Chinese and foreign theses?

(2) What are the similarities and differences in three elements including anaphors, antecedents, and distances used by native Chinese-speaking masters and native English-speaking masters?

(3) What are the underlying reasons for discrepancies in the total frequency and specific usage of third personal pronoun endophora applied by Chinese and foreign masters?

3.2. Research Instruments.

240 English abstracts of Chinese theses are selected from CNKI (China National Knowledge Infrastructure) and 240 English abstracts of foreign theses are selected from PQDT (ProQuest Dissertations & Theses) respectively. CNKI and PQDT are outstanding representatives of Chinese and foreign academic resource databases. AntConc software is applied to transform 480 English abstracts into a corpus. Those English abstracts with more than 100,000 words cover the disciplines of history, literature, sociology, and anthropology. With the aid of concordance, word list, file view, collocates, and other functions in AntConc software, relevant contexts of third personal pronoun anaphors are retrieved. On the basis of the "Hit" column on the left side of the software, the author annotated and analyzed three elements, including anaphors, antecedents, and distances.

3.3. Research Procedures.

Procedures of data collections consist of three steps:

(1) The first step is to automatically search "He," "His," "Him," "Himself," "She," "Her," "Herself," "It," "Its," "Itself," "They," "Their," "Them," and "Themselves" as third personal anaphors in the corpus.

(2) The second step is to artificially analyze the consistency of semantic components between current topics and third personal anaphors by applying the specific pragmatic context and the reader's cognitive capacity. Then the author marks the antecedent as "1" (1 means the antecedent is consistent with keywords of the thesis), and "2" (2 means the antecedent is inconsistent with keywords of the thesis).

(3) The last step is to measure the number of words between anaphors and antecedents and consider the distance as the third element, except for anaphors and antecedents.

4. Data Analysis and Discussion

4.1. Frequency of Third Personal Pronoun Endophora.

As endophora is divided into anaphora and cataphora, the total frequency of third-personal pronoun endophora in this research has been counted by the summation of anaphora and cataphora. In 240 English abstracts written by native Chinese-speaking masters, the total frequency of third personal pronoun endophora is 1329, in which there are 1258 anaphora examples and 71 cataphora examples. In 240 English abstracts written by native English-speaking masters, the total frequency of third personal pronoun endophora is 952, of which there are 899 anaphora examples and 53 cataphora examples. In order to test if there exists a significant difference in the frequency of third personal pronoun endophora used by Chinese and foreign masters, the author made use of the Chi-square test in that it mainly deals with the 2×2 table for testing the significant difference.

Because of $n > 40$ and an expected count > 1, the result of continuity correction[b] is emphasized in Table 1. The Sig. value of continuity correction[b] is 0.889, which is much larger than 0.05, hence there is no significant difference in the frequency of third personal pronoun endophora. In other words, both native Chinese-speaking masters and native English-speaking masters tend to apply third personal pronouns to refer to antecedents that appear in the previous or subsequent discourse.

In CNKI and PQDT, the frequency of third personal pronoun anaphora in the disciplines of history and literature accounts for the largest proportion. In CNKI, the cumulative percent in the two disciplines is 58.3%, and in PQDT, the cumulative percent is 65.9%. Masters who major in the disciplines of history and literature are inclined to employ the third personal pronoun endophora. The reason is that the discipline of history mostly involves historical figures, events, and places, and the discipline of literature frequently consists of characters, most of which are proper nouns. When Chinese and foreign masters refer to proper nouns that have appeared in English abstracts, they would like to make use of third personal pronouns to replace redundant antecedents. Only in this way, can the purpose of following the economic principle of language and guaranteeing discourse coherence be achieved.

Eg1. Based on close reading and text analyses of about 40 historical novels published around 1962, this thesis investigates the differences between these novels, as well as their values and limitations, and demonstrates the profound impact of two of Chen Xiang-he's novels on Chinese literature (sentence from the discipline of literature in CNKI).

Eg2. This thesis broadly analyzes the principles of Dante's aesthetic in the poem while analyzing the Sibyl, the Hanged Man, and the Prajapati for their relevance to Eliot's aesthetic theory (sentence from the discipline of literature in PQDT).

In two instances, the third personal pronoun "their" is the economic and succinct anaphor, but "40 historical novels published around 1962" and "the Sibyl, the Hanged Man, and the Prajapati" play the roles of complex and lengthy antecedents. Anaphors acted by third personal pronouns can not only make language succinct in form but also benefit readers in comprehension. That's why third personal pronoun endophora accounts for such a large proportion of the total, no matter in CNKI or PQDT. Both Chinese and foreign masters pay attention to the economic principle of language and focus on the least amount of effort in language and the greatest effect in communication. They tend to apply third personal pronouns in English abstracts so as to reduce complicated and redundant antecedents.

TABLE 1: Chi-square tests of third-personal pronoun endophora in CNKI and PQDT.

	Value	df	Asymp. Sig. (2-Sided)	Exact sig. (2-sided)	Exact sig. (1-sided)
Pearson Chi-square	0.055a	1	0.815		
Continuity correction[b]	0.020	1	0.889		
Likelihood ratio	0.054	1	0.816		
Fisher's exact test				0.852	.443
Linear-by-linear association	0.055	1	0.815		
N of valid cases	2281				

a. 0 cells (0.0%) have an expected count of less than 5. The minimum expected count is 51.75. b. Computed only for a 2×2 table.

With the aid of AntConc software and SPSS, 71 genuine cataphora cases were discovered in 240 English abstracts selected from CNKI, and 53 genuine cataphora cases were found in 240 English abstracts selected from PQDT. The author concludes that there exist two types of antecedents. The first type of antecedent is the infinitive form like "to describe..." and the second type of antecedent is the that-clause, like "that the history of Yingzhou Fu has a...". There is only one type of third personal anaphor. It is the third personal pronoun, "it". Summarized patterns and specific examples of cataphora in English abstracts written by native Chinese-speaking masters and native English-speaking masters are as follows:

It is +adj. (like impossible/necessary/...) + to do

Something makes it + adj. (like significant/important...) + to do

It is +v-ed (like argued/published/...) + that-clause

It is +adj. (like significant/important...) + that-clause

Something makes it + adj. (like significant/important...) + that-clause

Eg3. It is not normal to do research that associates preschool teachers with professional gender segregation (sentence from the discipline of sociology in CNKI).

Eg4. Through comparisons of the Broome County Department of Social Services end-of-year caseload reports of 2005 and 2010, it is clear that the number of caseloads has approximately doubled during this time period (sentence from the discipline of sociology in PQDT).

Eg5. Secondly, the disunity among Jewish organizations made it difficult to put them together as one voice to defend against the destruction of European Jews (sentence from the discipline of literature in CNKI).

Eg6. To sum up, a comparative study of these two books makes it clear that human beings can only build an ecological society by building a harmonious relationship with nature (sentence from the discipline of literature in PQDT).

After analyzing and discussing third personal pronoun anaphora and cataphora in English abstracts selected from CNKI and PQDT, it is concluded that the usage of the third personal pronoun endophora in English abstracts of theses should be encouraged. On the one hand, third personal pronouns, unlike first or second personal pronouns, are more objective, therefore they are in concordance with the objectivity of theses. On the other hand, the central idea of endophora is to use shorter language representations to substitute longer antecedents. Hence the phenomenon of endophora is the inevitable result of language objectivity and language economy.

4.2. Three Elements of Third Personal Pronoun Endophora. In order to test if there is a significant difference in the usage of three elements, including anaphors, antecedents, and distances, used by native Chinese-speaking masters and native English-speaking masters, the author performed an MNOVA test on three elements. The MNOVA test mainly deals with the relationship between two or more fixed factors and various dependent factors. However, according to statistical principles, before the MNOVA test is conducted, the test of homogeneity of variances has been accomplished. As the Sig. value of the test of homogeneity of variances is 0.000, which does not meet the standard of homoscedasticity, in the MNOVA test, "Tamhane's T2" below the line of equal variances not assumed is emphasized. Sig. values in Table 2 show that there exists a significant difference in the usage of anaphors, antecedents, and distances of third personal pronoun endophora used by Chinese and foreign masters.

In terms of anaphors used by Chinese masters, the anaphor "it" has the largest number of 97 in history, the largest number of 93 insociology, and the largest number of 87 in anthropology. The anaphor "it" used by native Chinese-speaking masters is frequently to describe some proper nouns. With regard to anaphors used by foreign masters, the anaphor "their" has the largest number of 70 in history, the largest number of 75 in sociology, and the largest number of 63 in anthropology. The anaphor "they" used by native English-speaking masters is frequently to depict some social spotlights. There are significant differences between the usage of anaphors in the discipline of literature and the usage of anaphors in the other three disciplines. In English abstracts written by Chinese masters, the frequency of the anaphor "it" is the highest in the disciplines of history, sociology, and anthropology, while the frequency of anaphor "his" is the highest in the discipline of literature. In English abstracts written by foreign masters, the highest frequency of anaphor in the disciplines of history, sociology, and anthropology is "they," while the highest frequency of anaphor in the discipline of literature is "her."

Table 2: Multivariate tests[b] of three elements of third personal pronoun endophora in CNKI and PQDT.

	Effect	Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's trace	0.962	13684.131[a]	3.000	2152.000	0.000
	Wilks' lambda	0.038	13684.131[a]	3.000	2152.000	0.000
	Hotelling's trace	25.435	13684.131[a]	3.000	2152.000	0.000
	Roy's largest root	25.435	13684.131[a]	3.000	2152.000	0.000
Source	Pillai's trace	0.066	38.061[a]	3.000	2152.000	0.000
	Wilks' lambda	0.934	38.061[a]	3.000	2152.000	0.000
	Hotelling's trace	0.071	38.061[a]	3.000	2152.000	0.000
	Roy's largest root	0.071	38.061[a]	3.000	2152.000	0.000

a. exact statistic and b. design: intercept + source.
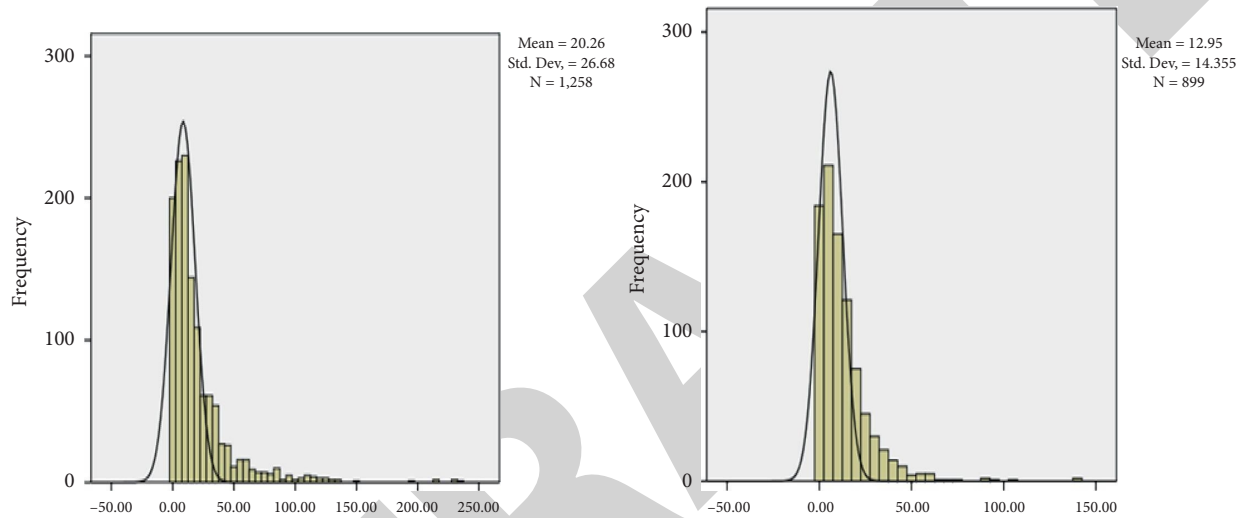
Figure 1: Distances between antecedents and anaphors in CNKI and PQDT.

It is discovered that Chinese and foreign masters majoring in the discipline of literature tend to prefer the third personal pronoun, "his" or "her," which is determined by professional characteristics. The discipline of literature mostly covers proper nouns for characters in works of literature. In order to avoid cumbersome repetitions of proper nouns, Chinese and foreign masters will use third personal pronouns. As the other three disciplines mostly involve proper nouns and social spotlights, Chinese and foreign masters would like to choose "it" or "their" to replace antecedents in English abstracts for achieving the objective purpose.

At the mention of antecedents applied by native Chinese-speaking masters and native English-speaking masters, there are significant differences in the usage of antecedents among disciplines. In English abstracts, antecedents used by Chinese masters are mostly consistent with keywords of the thesis, but antecedents used by foreign masters are mostly inconsistent with keywords of the thesis. Compared with native Chinese-speaking masters, native English-speaking masters are more inclined to direct readers to experience the whole theme of the thesis rather than simply highlight the topic of the thesis by using repetitions. This difference is traced back to discrepant mindsets and is consistent with the research conducted by Frauen and Shuneng LIAN.

According to two scholars, the mindset of Chinese masters displays holistic and inductive characteristics, while the mindset of foreign masters shows analytic and deductive features [16, 17]. The fundamental difference between Chinese and foreign mindsets lies in the perceptual thinking model and the reasonable thinking model. The former employs images to organize perceptual materials so that integral characteristics are portrayed. The latter makes use of logic and reasonable thinking to explore the essence and internal connections of materials so that analytical characteristics are depicted. The phenomenon that native Chinese-speaking masters frequently repeat keywords of the thesis reflects the intuition-oriented mindset, highlighting the whole-to-part pattern and the inductive thinking model. The phenomenon that native English-speaking masters rarely repeat keywords of the thesis embodies the experience-oriented mindset, highlighting the part-to-whole pattern and the deductive thinking model.

With regard to distances between antecedents and anaphors set by native Chinese-speaking masters and native English-speaking masters, there are significant differences in the usage of distances among the four disciplines. No matter the Kurtosis or Skewness of distances, both values exceed the normal distribution value of 3 or 0. In statistics, if values of Kurtosis and Skewness are larger than normal distribution

values, the histogram of distances definitely shows the features of the aiguille and fat tail branch.

According to Figure 1, there exist centered frequency distributions and large ranges of distance between antecedents and anaphors. Specifically, the number of words set between antecedents and anaphors by Chinese and foreign masters is controlled within a range of 0–50 words, but the average distance between antecedents and anaphors set by Chinese masters is 20.26 words, which is much longer than that set by foreign masters. The average distance set by foreign masters is merely 12.95 words. If the distance of 12.95 words is regarded as the distance threshold between antecedents and anaphors, the average distance set by Chinese masters exceeds this threshold.

5. Conclusion

The corpus is applied to compare and analyze the third personal pronoun endophora and its three elements in English abstracts of Chinese and foreign theses. Underlying reasons for similarities and differences are explored in order to help native Chinese-speaking masters have a better understanding of international writing conventions of English abstracts and to further improve the language quality of English abstracts in their theses.

In 480 English abstracts, there is no significant difference in the total frequency of third personal pronoun endophora used by native Chinese-speaking masters and native English-speaking masters. Driven by the economic principle of language, Chinese and foreign masters tend to avoid redundant words, reduce the memory burden of encoding and decoding, and finally improve the efficiency of language communication. Chinese masters pay attention to the cognitive schema from the whole to the part, and therefore the antecedents they use are usually consistent with keywords of the thesis. The direct repetition not only achieves the overall consistency of the thesis and the prominence of the theme, but also demonstrates the integrity and induction of the mindset. On the contrary, foreign masters show solicitude for the cognitive schema from the part to the whole, and hence the antecedents they apply are hardly consistent with the keywords of the thesis. The indirect description can not only lead readers to explore the whole topic when they experience different parts but also reflect the analytical and deductive mindset. Although the distance between antecedents and anaphors set by Chinese and foreign masters displays the same feature as the aiguille and fat tail branch, the distance set by Chinese masters far exceeds the threshold set by foreign masters. This finding has some implications for Chinese masters. When they use third personal pronoun endophora in English writing, they are suggested to shorten the distance between antecedents and anaphors in order to achieve discourse coherence.

It is concluded that native Chinese-speaking masters had better make efforts in terms of language representation and deep cognition in order to get closer to the international English abstract writing conventions. With regard to the language representation, Chinese masters are suggested to reduce the distance between antecedents and anaphors

because the excessive distance will increase the memory burden of encoding and decoding for readers and have an influence on the readability and coherence of the discourse. In terms of deep cognition, 30 Chinese masters interviewed realize that they should pay attention to the cognitive schema of exploration and analysis and imitate the mindset of foreign masters. To be specific, the repetitiveness relationship between antecedents and keywords should be adjusted to the relevance relationship in order to comply with foreign masters' English abstract writing convention. Only the language representation and deep cognition are concerned, the language quality of English abstracts in their theses will be ultimately enhanced.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] M. A. K. Halliday and Hasan, *Cohesion in English*, Foreign Language Teaching and Research Press, Beijing, 2001.

[2] H. Lu, "A quantitative and comparative study on the usage of personal pronouns in college English writing," *Henan Normal University*, vol. 34, no. 4, pp. 157-158, 2007.

[3] H. Helen, N. Tiina, and P. Renate, "Pronouns as referential devices in Estonian, Finnish, and Russian," *Pragmatics*, vol. 155, pp. 43–63, 2020.

[4] X. U. Jiangli, *Error Analysis and Teaching Suggestions of Anaphora in the Writing of English Speaking Advanced Chinese Learners*, Dalian University of Foreign Language, Dalian, China, 2021.

[5] M. Strumpf, *The Grammar Bible*, Henry Holt, New York, NY, USA, 2004.

[6] H. Wang, *A Study of the Evolution of English Third Person Pronouns from the Perspective of Cognitive Sociolinguistics*, Sichuan International Studies University, Chongqing, China, 2021.

[7] N. Chomsky, *Lectures on Government and Binding*, Foris Publications, Dordrecht, 1981.

[8] S. Levinson, *Pragmatics*, Cambridge University Press, Cambridge, 1983.

[9] K. Fukumura and R. P. G. van Gompel, "Choosing anaphoric expressions: do people take into account likelihood of reference?" *Journal of Memory and Language*, vol. 62, no. 1, pp. 52–66, 2010.

[10] H. G. Widdowson, *Discourse Analysis*, Shanghai Foreign Language Education Press, Shanghai, China, 2012.

Research Article

FoSSA Optimization-Based SVM Classifier for the Recognition of Partial Discharge Patterns in HV Cables

Kang Sun [1,2] **Yuxuan Meng** [1] **and Shuchun Dong** [2]

[1]*School of Electrical Engineering and Automation,*
 Henan Key Laboratory of Intelligent Detection and Control of Coal Mine Equipment, Henan Polytechnic University,
 Jiaozuo 454003, China
[2]*Dianrong Intelligent Technology Co., Ltd., Kunshan 215334, China*

Correspondence should be addressed to Kang Sun; sunkang@hpu.edu.cn

In order to enhance the classification accuracy and the generalization performance of the SVM classifier in cable partial discharge (PD) pattern recognition, a firefly optimized sparrow search algorithm (FoSSA) is proposed to optimize its kernel function parameters and penalty factors. First, the Circle-Gauss hybrid mapping model is employed in the population initialization stage of the sparrow search algorithm (SSA) to eliminate the uneven population distribution of random mapping. Sparrows tend to fall into local extremums during the search process, while the firefly algorithm has a fast optimization speed and strong local search ability. Thus, a firefly disturbance is added in the sparrow search process, and the fitness value is recalculated to update the sparrow position to enhance the sparrow's local optimization ability and accuracy. Finally, based on the SSA, a dynamic step-size strategy is adopted to make the step size dynamically decrease with the number of iterations and improve the accuracy of convergence. Six benchmark functions are employed to evaluate the optimization performance of the FoSSA quantitatively. Experiment results show that the recognition accuracy of the PD patterns using the SVM optimized by the FoSSA could reach 97.5%.

1. Introduction

Power cable is the key infrastructure equipment for urban distribution networks and large-scale clean energy access, and its reliability is vital to the safe and stable operation of the power system [1]. The defects of insulation material, manufacturing process, and structural defects, coupled with the aging of insulation material caused by harsh electrical, thermal, and mechanical stresses environment, will result in a partial discharge (PD) and even dielectric breakdown, which lead to insulation failure [2]. Accurate and instant identification of the fault pattern by mining and analyzing operation fault records and all types of test data of the power cable can significantly improve the efficiency for the maintenance and overhaul of a cable system.

Due to the low frequency of faults during operation, imperfect records of fault, and abnormal information, the size of fault samples is usually limited. Benefitting from the

structural risk minimization (SRM) criteria and the kernel methods, support vector machines (SVMs) [3] have shown significant superiority to deal with the classification problems of few samples and nonlinear high-dimensional data. Thus, it has been widely applied in the fault pattern recognition of large-scale electrical equipment, such as cables, transformers, and power grids [4–7]. However, the classification performance of the SVM is highly dependent on the selection of kernel function parameters and penalty factors, so how to optimize the parameters is crucial for its further applications.

For deterministic optimization algorithms, such as sequential minimal optimization (SMO) [8] and stochastic gradient descent (SGD) [9], if the objective function is discontinuous and nondifferentiable, their convergence speed is usually slow and they will easily fall into the local optimum. As a stochastic optimization method, the swarm intelligence optimization methods introduce a brand new

path to solve global optimization problems by taking advantage of randomness. The particle swarm optimization (PSO) algorithm [10] and the ant colony optimization (ACO) algorithm [11] are the most representative of these.

The PSO algorithm has few parameters and a fast convergence speed, but it tends to fall into local extremes due to premature convergence. This can be improved partly by introducing inertial weighting factors and taboo detection mechanisms. However, for complex high-dimensional problems, usually, it is impossible to guarantee convergence to the global optimum. The ACO algorithm uses the positive feedback mechanism of ant colony pheromones to strengthen the learning ability. Its heuristic probabilistic search mode makes it not easy to trap in the local optimum. However, the parameter settings are complicated and searching speed is slow; furthermore, the convergence property is pretty poor. In order to further strike a balance between the search range and the convergence accuracy in optimization algorithms, a series of bionic intelligent optimization algorithms, such as gray wolf optimization algorithm [12] (GWO), artificial bee colony algorithm [13] (ABC), and the bacterial foraging algorithm [14] (BFA), have been proposed in recent years. Sparrow search algorithm (SSA) [15] is a novel swarm intelligence optimization algorithm that is inspired by foraging and antipredation behaviors of sparrows. Testing results on unimodal and multimodal functions demonstrate its superiority over PSO, ACO, and GWO in terms of accuracy, convergence speed, stability, and robustness.

This paper proposes a FoSSA-optimized SVM for the recognition of partial discharge patterns in HV cables. First, the feature vector is constructed based on the partial discharge $\varphi - q - n$ pattern. Second, in the standard SSA, the Circle-Gauss hybrid mapping model is introduced to initialize the population to improve the diversity. In the sparrow search process, the dynamic step strategy and the firefly interference strategy are introduced to make the sparrow escape from the local optimum and find the global optimal combination of support vector machine parameters. Finally, an optimized SVM classification framework is constructed for the partial discharge recognition in the HV power cables.

2. Problem Description

The basic principle of a nonlinear SVM is to map the input space x to a feature space $\Phi(x)$ through a nonlinear transformation which results in a hyperplane model in the feature space corresponding to the hypersurface model in the input space. The hyperplane in the feature space is divided as follows:

$$f(x) = \omega^T \Phi(x) + b, \tag{1}$$

where ω is the weighting vector and b is the threshold.

For a given nonlinear separable data set, considering the existence of errors ξ, the optimization problem with the constraint conditions is as follows:

$$\min \frac{1}{2} \|\overrightarrow{\omega}\|^T \omega + c \sum_{i=1}^{l} \xi_i, \tag{2}$$

$$\text{st.} y_i \left(\overrightarrow{\omega}^T \Phi(x_i) - b \right) \geq 1 - \xi_i (\xi_i > 0).$$

The optimization of equation (2) can be transformed into a dual problem by introducing the Lagrange factor, and the solution of equation (1) can be obtained as follows:

$$f(x) = \sum_{i=1}^{l} \alpha_i \kappa(x, x_i) + b, \tag{3}$$

where α_i is the Lagrange factor, l is the number of support vectors, and $\kappa(x, x_i)$ is the kernel function. A radial basis function (RBF) in equation (4) is generally adopted.

$$\kappa(x, x_i) = \exp \left(-g \|x - x_i\|^2 \right). \tag{4}$$

It can be seen from the derivation that the parameter selection of c and g directly affects the classification performance of SVM. In the traditional SVM model, they are usually selected according to expert experience or k-fold cross verification. In the process of cable fault classification, the input data are diverse and complex. The parameter selection based on experience not only takes time but also brings some randomness to the calculation process. The K-fold cross-verification method is dependent on the parameter range; if this range is inappropriate, it is impossible to determine the optimal parameters.

3. Firefly Optimized Sparrow Search Algorithm (FoSSA)

3.1. SSA Principle. SSA is a novel swarm intelligence algorithm that has evolved from the foraging and antipredation behaviors of sparrows. The algorithm is simple and efficient, and it can achieve global convergence. According to the mathematical model of the algorithm, virtual sparrows are used for foraging behavior and the position of sparrows can be expressed as follows:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix}, \tag{5}$$

where d is the dimension of the variable in the optimization question, n is the number of sparrows, and $x_{i,j}$ is the position of the i-th sparrow of the j-th dimension.

The fitness value of all sparrows can be calculated as the following vector:

$$\mathbf{F}_x = \begin{bmatrix} f(x_{1,1}, x_{1,2}, ..., x_{1,d}) \\ f(x_{2,1}, x_{2,2}, ..., x_{2,d}) \\ \vdots \\ f(x_{n,1}, x_{n,2}, ..., x_{n,d}) \end{bmatrix}, \tag{6}$$

where f represents the individual fitness value.

A sparrow population can be divided into producer and scrounger according to the relative role of each sparrow. Producers are in charge of looking for food for the whole population; they provide foraging directions for the scroungers. Producers can obtain a larger foraging area than the scroungers. According to the foraging rules of sparrow population, the moving position of the producer is calculated as follows:

$$\mathbf{X}_i^{t+1} = \begin{cases} \mathbf{X}_i^t \cdot e^{-\left(i/\alpha.t_{\max}\right)} & \text{if } V_a < V_s, \\ \mathbf{X}_i^t + r_d \cdot \mathbf{U} & \text{if } V_a \geq V_s, \end{cases} \quad (7)$$

where $\mathbf{X}_i^t$ is the position of the i-th sparrow in j-dimension space at iteration t, $i \in [1, 2, \dots, n]$, t indicates the current iteration, $t_{\max}$ is the maximum iteration number, $\alpha \in (0, 1]$ is a random number, r_d is a random number that obeys a normal distribution, U represents a unit matrix of $1 \times d$, and $V_a \in [0, 1]$ and $V_s \in [0.5, 1]$ are the alarm and safety value relative to the predators, respectively, and they determine the sparrow's moving range.

During the foraging process, the scroungers keep eyes on the producers. Once the producers find something better, they fight for it immediately. According to the rules, their moving position can be updated as follows:

$$\mathbf{X}_i^t = \begin{cases} r_d \cdot e^{\left(\mathbf{X}_{wst}^t - \mathbf{X}_i^t\right)/i^2} & \text{if } i > n/2, \\ \mathbf{X}_{op}^{t+1} + \left\| \mathbf{X}_i - \mathbf{X}_{op}^{t+1} \right\| \cdot \mathbf{A}^+ \cdot \mathbf{U} & \text{others,} \end{cases} \quad (8)$$

where $\mathbf{X}_{op}$ is the optimal location of the producers, $\mathbf{X}_{wst}^t$ is the current global worst location, $\mathbf{A}$ is the matrix of $1 \times d$ with elements are all 1 or -1 randomly, and $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$.

In the search process, some sparrows called guards will be aware of the danger from the predators, according to the antipredation rule, and the mathematical expression of their moving positions can be obtained as follows:

$$\mathbf{X}_i^{t+1} = \begin{cases} \mathbf{X}_{bst}^t + \beta.\left\| \mathbf{X}_i^t - \mathbf{X}_{bst}^t \right\| & \text{if } f_i > f_b, \\ \mathbf{X}_i^t + \gamma.\left(\dfrac{\left\| \mathbf{X}_i^t - \mathbf{X}_{wst}^t \right\|}{\left((f_i - f_w) + \varepsilon\right)} \right) & \text{if } f_i = f_b, \end{cases} \quad (9)$$

where $\mathbf{X}_{bst}^t$ is the current global optimal value, β is a random parameter obeying standard normal distribution that constraints step size, $\gamma \in [-1, 1]$ is a random number, f_i is the fitness value of the current sparrow, f_b and f_w are the best and worst fitness value, and ε is a regulatory factor.

3.2. Initialization by Circle-Gauss Hybrid Mapping.

The distribution of the initial population is important for SSA. A uniform and fully mapped initial distribution will effectively improve the convergence speed of the optimization process. Due to the lack of initialization strategy for uniform population distribution in SSA, simple random distribution cannot guarantee the breadth of the search range, and it is easy to produce "super sparrows" in the iterative process that

cause other individuals to gather to them, resulting in a "premature" phenomenon and reducing the diversity of the population.

In this study, a Circle-Gauss hybrid mapping model is introduced to initialize the SSA. By combining the advantage of the regularity and uniformity of Circle mapping and the randomness and ergodicity of Gauss mapping, the chaotic sequence can be transformed into the solution space of the SSA algorithm to replace the original population by Circle-Gauss hybrid mapping model. The Circle-Gauss hybrid mapping model not only avoids the overdensity of the population but also retains the diversity of the population to a large extent, which is the key factor for the global optimization of the SSA algorithm. The mathematical expression of the Circle-Gauss hybrid mapping model is as follows:

$$\mathbf{M}_{i+1} = \begin{cases} \mathbf{M}_i + a - \text{mod}\left(\dfrac{b}{2\pi} \sin\left(2\pi\mathbf{M}_i\right), 1 \right) & \text{if } i = 2k; \\ \\ \dfrac{1}{\mathbf{M}_i} - \left[\dfrac{1}{\mathbf{M}_i} \right] & \text{if } i = 2k+1; \end{cases} \quad (10)$$

where $\mathbf{M}_i \in \mathbb{R}^{1 \times d}$ is the mapping position of i-th sparrow, $\mathbf{M} = [\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_n]^T \in \mathbb{R}^{n \times d}$ is the mapping position of the whole population, n and d have the same meaning as in equation (5). a is 0.5, b is 2.2, and $k = 0, 1, 2\dots$. mod() and [] represent modulation and rounding operation.

Figure 1 shows the 2D scatter diagram generated by the Circle-Gauss hybrid mapping model and the other two mappings in (0, 1). As shown in Figure 1, the Circle-Gauss hybrid mapping model combines the characteristics of both uniformity and randomness, improves the ergodicity and effectiveness of the initialization, and ensures the diversity of the population.

3.3. Firefly Perturbation.

Firefly perturbation is introduced in SSA here to improve its global convergence ability while the sparrow falls into the local extremum. In the two-dimensional solution space of SVM parameter combination (c, g), analogous to the idea of the firefly algorithm, each sparrow is initialized by the attractiveness of ρ_0. During the disturbance process, the degree of attraction decreases with the increase of the spatial distance. When the sparrow traps in a local optimum, the perturbation will be introduced as follows:

(1) The spatial distance r between the sparrow position $\mathbf{X}_i^t$ and the optimal position $\mathbf{X}_{bst}^t$ is calculated as follows:

$$r = \frac{\sqrt{d_m}}{b_l - b_u} \left\| \mathbf{X}_i^t - \mathbf{X}_{bst}^t \right\|. \quad (11)$$

In the expression, b_l and b_u represent the lower and upper boundary, respectively, and d_m is the space dimension.
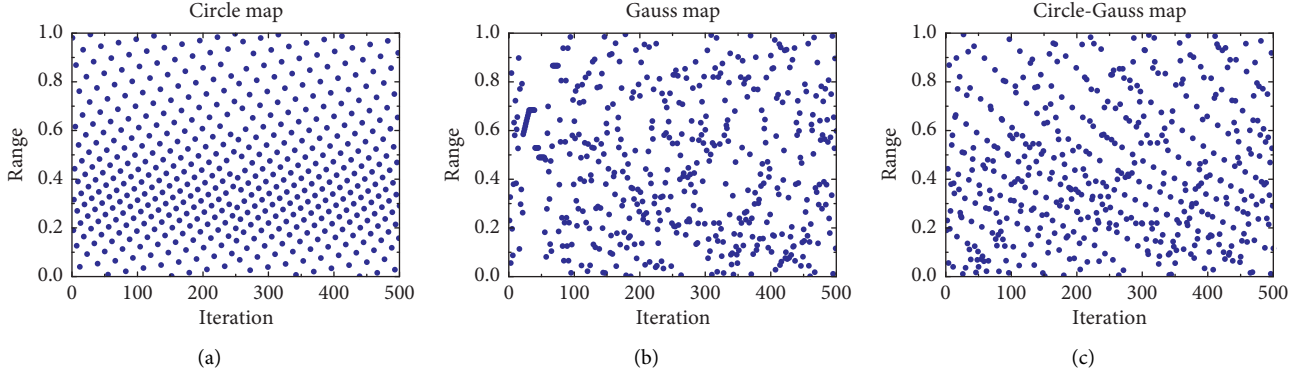
Figure 1: Population initialization scatter plot of three methods. (a) Circle map. (b) Gauss map. (c) Circle-Gauss map.

Table 1: Expressions of characteristic parameters.

Symbol of parameters	Meaning
sk_m^+	$\varphi-q_{\max}$ pos(+) half-cycle skewness
sk_m^-	$\varphi-q_{\max}$ neg(−) half-cycle skewness
sk_m	$\varphi-q_{\max}$ full-cycle skewnesss
sk_n^+	$\varphi-q_{\mathrm{mean}}$ pos(+) half-cycle skewness
sk_n^-	$\varphi-q_{\mathrm{mean}}$ neg(−) half-cycle skewness
sk_n	$\varphi-q_{\mathrm{mean}}$ full-cycle skewness
kt_m^+	$\varphi-q_{\max}$ pos(+) half-cycle kurtosis
kt_m^-	$\varphi-q_{\max}$ neg(−) half-cycle kurtosis
kt_m	$\varphi-q_{\max}$ full-cycle kurtosis
kt_n^+	$\varphi-q_{\mathrm{mean}}$ pos(+) half-cycle kurtosis
kt_n^-	$\varphi-q_{\mathrm{mean}}$ neg(−) half-cycle kurtosis
kt_n	$\varphi-q_{\mathrm{mean}}$ full-cycle kurtosis
r_m	$\varphi-q_{\max}$ ratio of the sum of pos(+) and neg(−) half-cycle discharge amplitudes
r_n	$\varphi-q_{\mathrm{mean}}$ ratio of the sum of pos(+) and neg(−) half-cycle discharge amplitudes

(2) The attractiveness value ρ of each individual sparrow is calculated:

$$\rho = \rho_0 e^{-\theta r^2}. \tag{12}$$

In the expression, ρ_0 is the maximum attractiveness and θ is the attractiveness coefficient.

(3) The sparrow position with perturbations $\mathbf{X}_i^t$ can be obtained as follows:

$$\mathbf{X}_i^{t'} = \mathbf{X}_i^t + \rho\left(\mathbf{X}_i^t - \mathbf{X}_{bst}^t\right) + \tau\left(\mathbf{R}_d - 0.5\right). \tag{13}$$

Here, τ is the step-size factor and $\mathbf{R}_d \in \mathbb{R}^{1 \times d}$ is a matrix with all elements obeying uniform distribution within $[0, 1]$.

3.4. Dynamic Step-Size Updating.
In the SSA, the step-size control parameters are constant, which cannot make the SSA achieve a balance between the local optimization and the global optimization in the iterative process. It will affect the effect and speed of the optimization.

In the initialization process, a longer step-size factor is adopted to enhance the algorithm's global search capability. In the later process, a smaller step is adopted to enhance the local optimization capability. Therefore, in this study, a dynamic updating strategy of step size is adopted to make

the step size decrease nonlinearly with the increase of iteration number for the guards in equation (9). The dynamic step size can be optimized as follows:

$$\lambda(t+1) = i_{\mathrm{m}}\lambda(t)e^{t/i_m}\left(\frac{\lambda_0}{\ln t} + r_d\left(1 - \lambda_0\right)\right), \tag{14}$$

where λ_0 is the initial step-size factor and i_m is the maximum iteration number.

4. Cable PD Pattern Recognition Based on FoSSA-SVM

4.1. Feature Vector Extraction.
Phase resolved partial discharge (PRPD) spectrum reveals the relationship between the number of PD signals with different peak values and the phase angle. Since the PRPD spectra of different defects achieve different distribution characteristics, statistical characteristic parameters based on PRPD can be used to recognized different insulation faults. In this study, 14 statistical characteristic parameters are employed for the feature extraction of the PD data. The expressions and the meanings of the characteristic parameters are shown in Table 1.

According to the calculation of statistical characteristic parameters, 9 key features are extracted as PD patterns, as shown in Table 2.

TABLE 2: Symbol definition of characteristic quantity.

Characteristic	Symbol description
Mean	μ
Variance	σ^2
Skewness	S_k
Steepness	K_u
Local peak number	P_e
Discharge factor	Q
Degree of phase asymmetry	$\varnothing$
Cross-correlation coefficient	C_c
Corrected cross-correlation coefficient	m_c
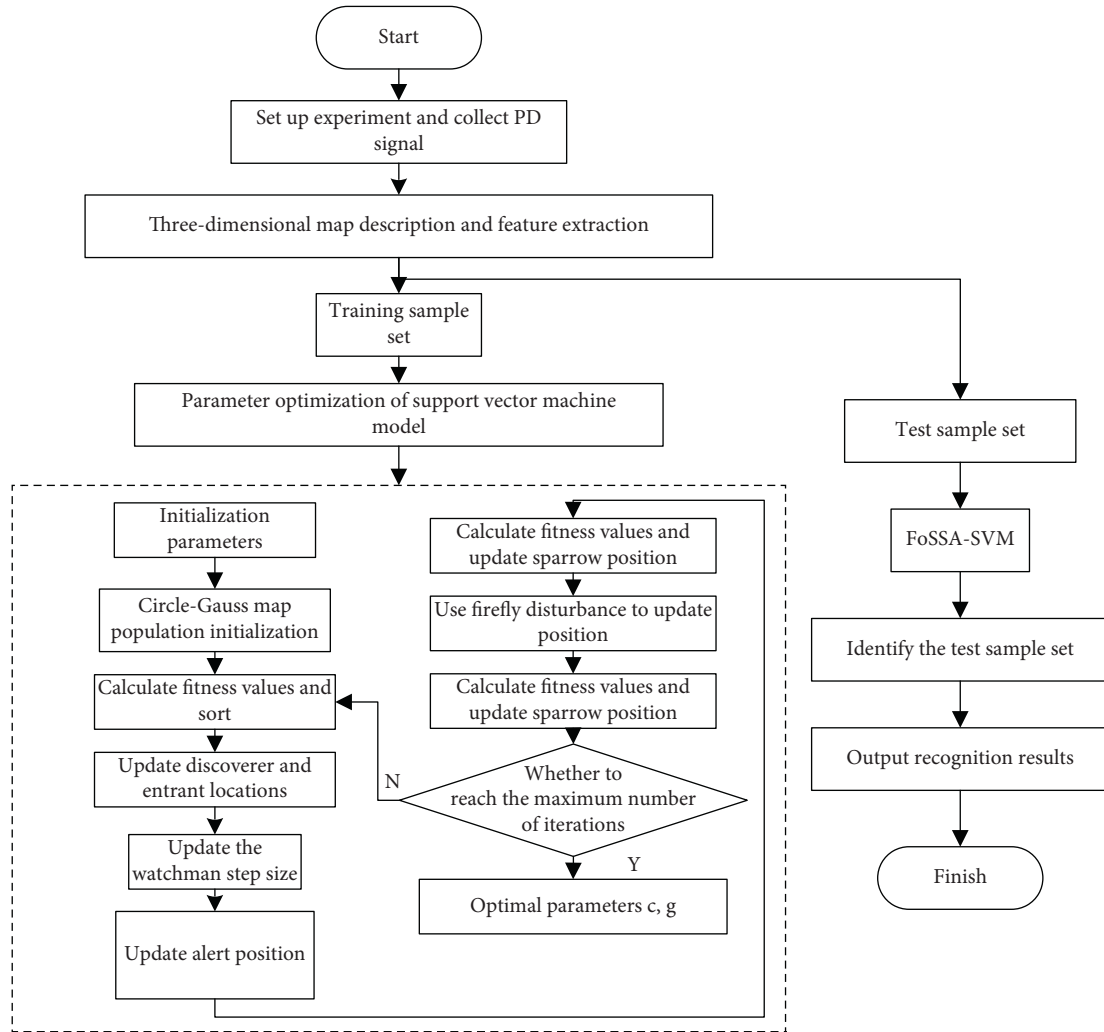
FIGURE 2: Flow chart of PD pattern recognition based on FoSSA-SVM.

The skewness S_k reflects the skewness of the spectrum shape compared with the normal distribution. The steepness K_u is used to describe the protrusion degree of the distribution of a shape compared with the normal distribution shape. The factor Q reflects the difference of the average discharge in the pos(+) and neg(-) half-cycle of the $\varphi - q$ spectrum. The phase asymmetry degree $\varnothing$ reflects the difference of the initial phase of the discharge within different half-cycles of the $\varphi - q$ spectrum. The cross-correlation coefficient C_c reflects the degree of shape similarity of the spectra in different half-cycles.

4.2. PD Pattern Recognition Based on the FoSSA-SVM Model. The FoSSA-optimized SVM model for cable PD pattern recognition proposed in this paper is shown in Figure 2. The whole scheme can be divided into feature extraction, parameter optimization, and pattern recognition:

(1) Four kinds of PD defects are manually prepared for data acquisition on a test platform. Then, the 3D PRPD spectrum is drawn for key feature extraction using statistical characteristics.

TABLE 3: Discharge voltage of different defects.

Defect type	Discharge voltage
Outer semiconductive layer creepage	5.6 kV, 6.6 kV
Internal air gap	5.3 kV, 11.3 kV, 18.3 kV
Scratch of insulation surface	5.6 kV, 9.6 kV, 13.6 kV
Metallic filth on insulation surface	6 kV, 20 kV, 34 kV

(a)

(b)

(c)

(d)
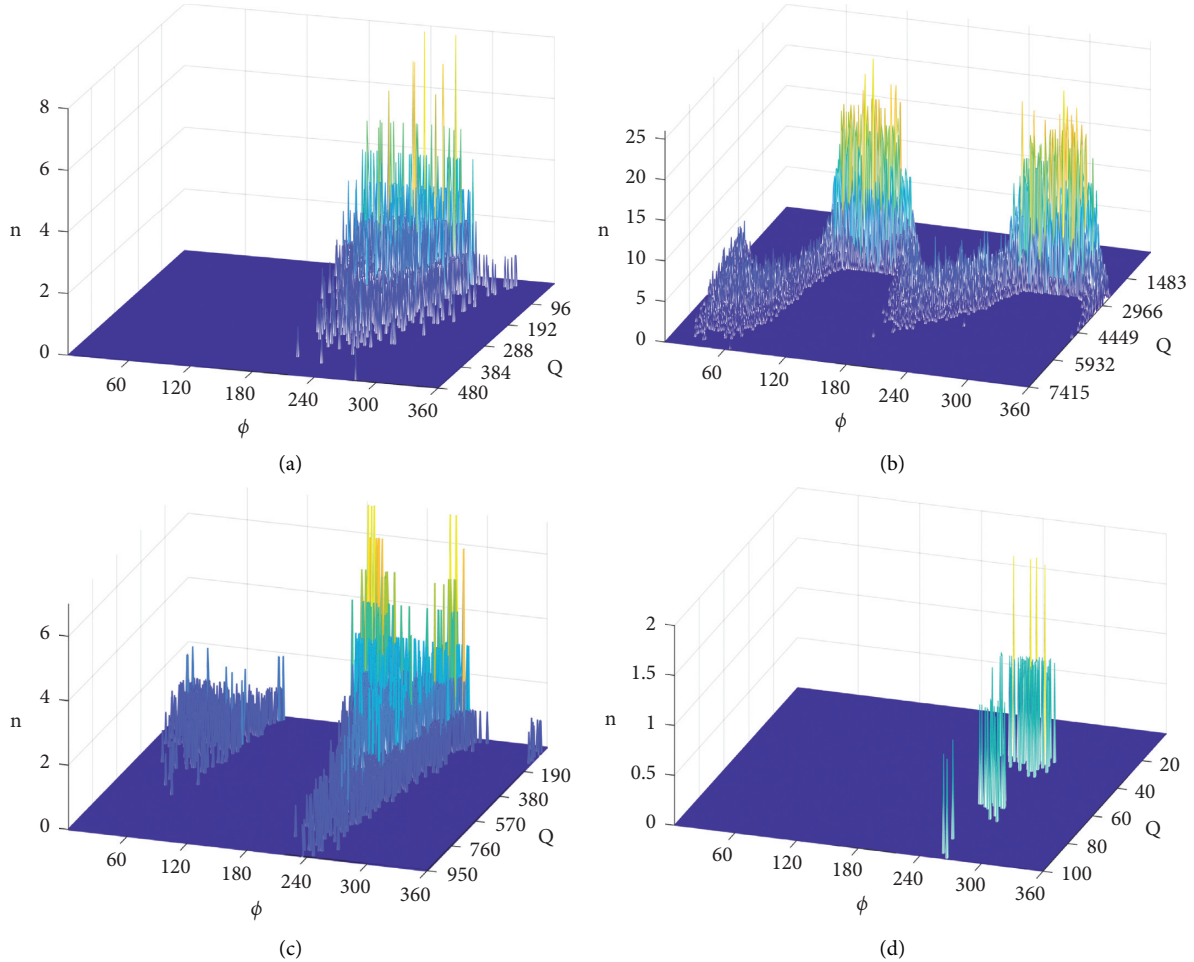
FIGURE 3: PRPD spectra of four type defects. (a) Outer semiconductive layer creepage (5.6 kV). (b) Internal air gap (5.3 kV). (c) Scratch of the insulation surface (5.6 kV). (d) Metallic filth (6 kV).

(2) The parameters of FoSSA and SVM are initialized, and the penalty factor and kernel function parameters of SVM are taken as the optimization objective of FoSSA. During the initialization stage, the Circle-Gauss hybrid mapping model is employed to initialize the population of sparrows.

(3) FoSSA algorithm is used for parameter optimization. The recognition error rate is set as the objective function for iterative calculation. The optimal parameter combination obtained is imported into the SVM model and verified by the test set. The detailed steps of the FoSSA to optimize the combination parameters are as follows:

Step 1: the size of sparrow population, number of iterations, producer and scrounger ratios, step-size parameters, and SVM parameters are initialized.

Step 2: the Circle-Gauss hybrid mapping model is used to generate the initial population of sparrows.

Step 3: the fitness value is computed and sorted.

Step 4: the positions of producers and scroungers are updated according to equation (7) and equation (8), respectively.

Step 5: the step size of the guards is updated according to equation (14).

Step 6: the position of the guards is updated according to equation (9).

TABLE 4: Test function expressions and their range.

Test function	Range
$f_1(x) = \sum_{i=1}^{n-1} \left[100\left(x_{i+1} - x_i^2\right)^2 + \left(x_i - 1\right)^2 \right]$	$[-30, 30]$
$f_2(x) = \sum_{i=1}^{n} i x_i^4 + \text{random}[0, 1)$	$[-1.28, 1.28]$
$f_3(x) = -20 \exp\left(-0.2\sqrt{1/n \sum_{i=1}^{n} x_i^2}\right) - \exp\left(1/n \sum_{i=1}^{n} \cos(2\pi x_i)\right) + 20 + e$	$[-32, 32]$
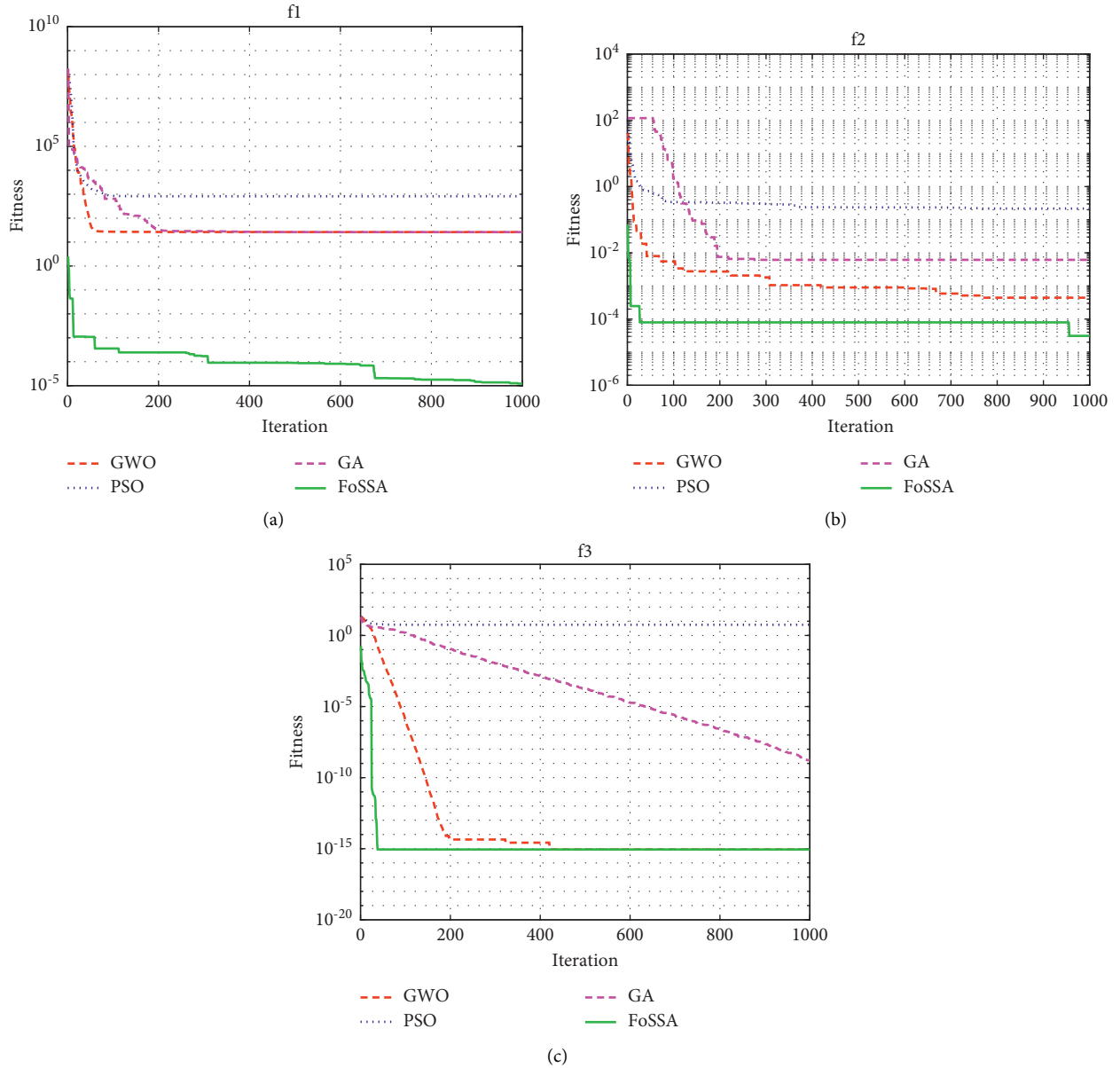
FIGURE 4: Convergence results of test functions. (a) Convergence results of $f_1(x)$. (b) Convergence results of $f_2(x)$. (c) Convergence results of $f_3(x)$.

Step 7: update the sparrow positions taking consideration of firefly disturbance according to equation (13).

Step 8: the fitness value is recalculated and the sparrow's position is updated.

TABLE 5: Function expressions and their range.

Test function	Range
$f_4(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{|x_i|}\right)$	$[-500, 500]$
$f_5(x) = \sum_{i=1}^{n} -x_i \cos\left(\sqrt{|x_i|}\right)$	$[-500, 500]$
$f_6(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{6} a_{ij}\left(x_j - p_{ij}\right)\right)^2$	$[0, 1]$

(a)

(b)

FIGURE 5: Continued.

Parameter space

Objective space

(c)
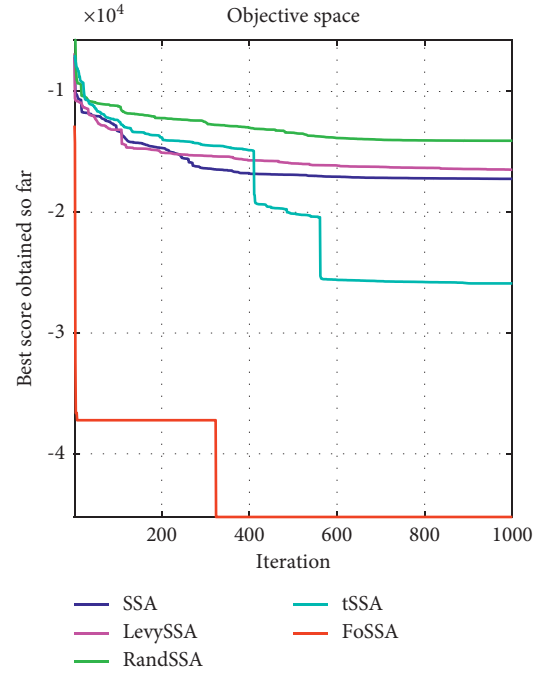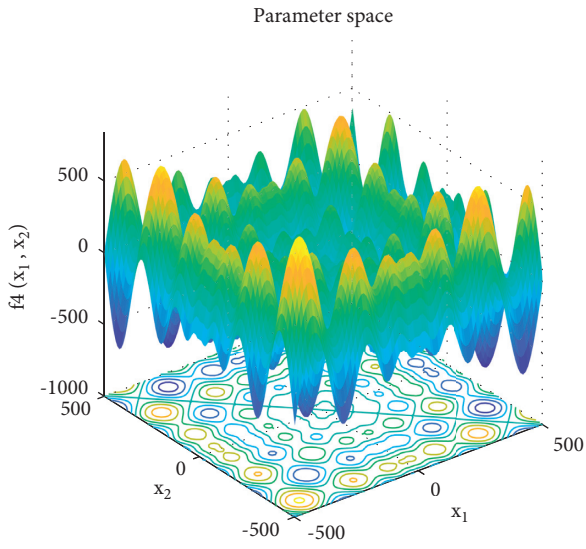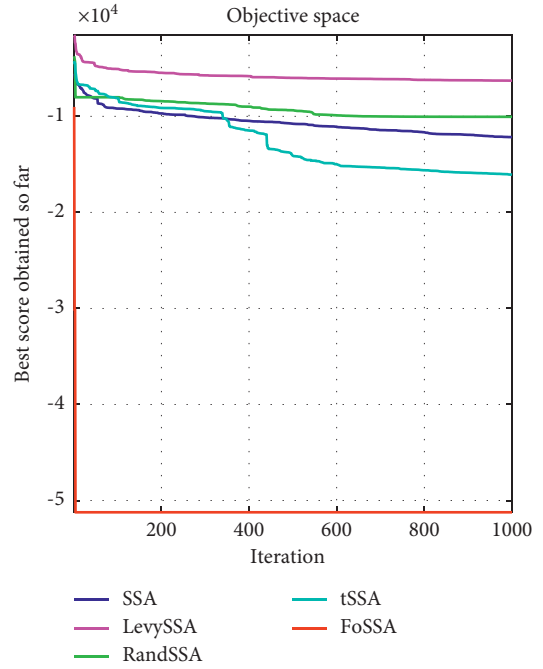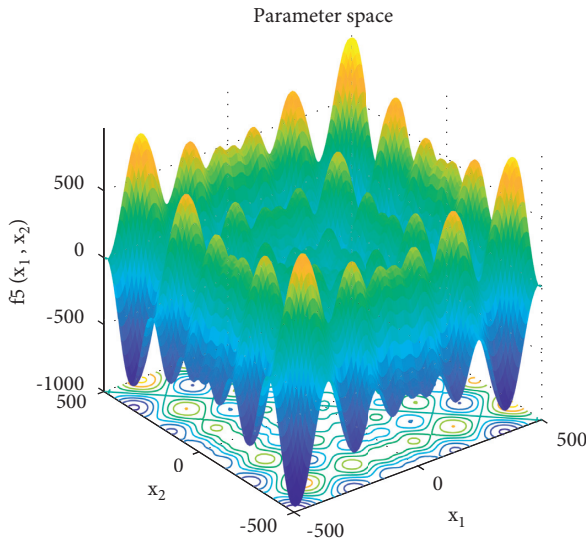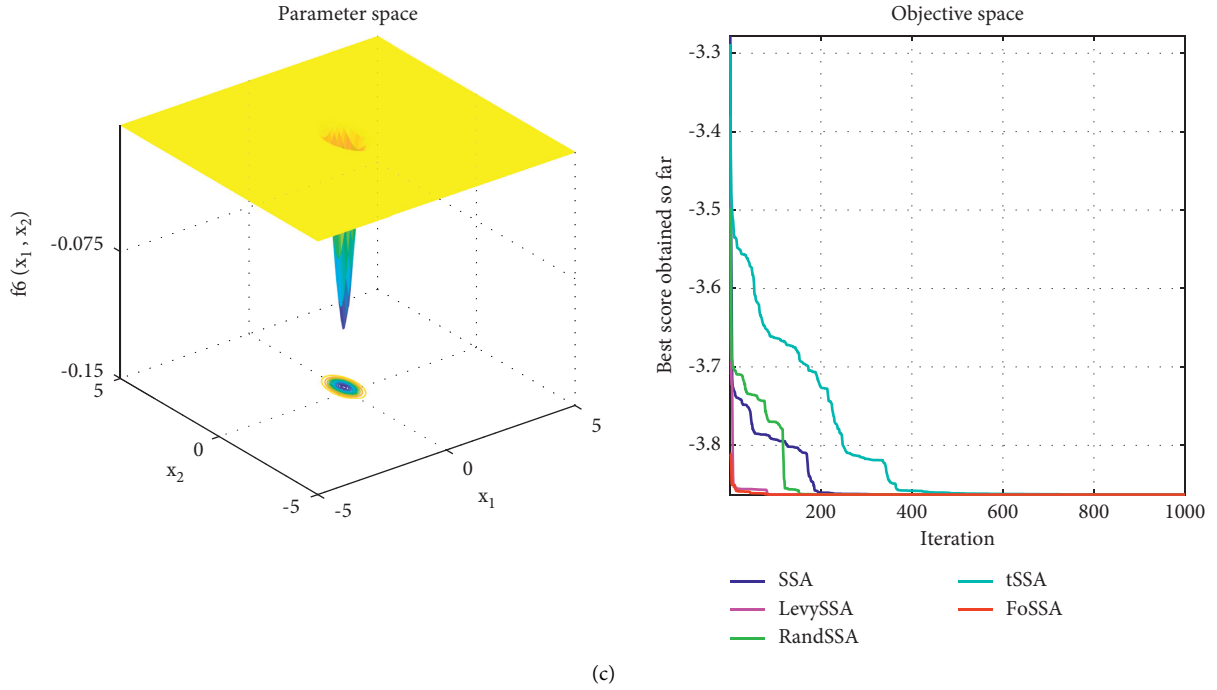
FIGURE 5: Convergence rate and optimal contour of test functions. (a) Convergence rate and optimal contour of $f_4(x)$. (b) Convergence rate and optimal contour of $f_5(x)$. (c) Convergence rate and optimal contour of $f_6(x)$.

Step 9: if the stop conditions are met, the algorithm is exited and the results are output. Otherwise, repeat Steps 3–8.

5. Experimental Results and Analysis

5.1. PD Sample Set Construction. According to the designed experiment, the cable PD data are acquired. The sampling rate of the oscilloscope used in the experiment is set at 10 MS/s, and the sampling length is 1s each time; that is, the signal containing 50 power frequency cycles is taken as one sample. The number of samples from each type of defect is 50; that is, each PD defect contains 50 samples. The voltage levels corresponding to the measurement of the four PD defects are shown in Table 3. After the collection of sampling points, the PRPD spectra with four defects are drawn, as shown in Figure 3. The training set is 80% of these defect samples and the test set is 20%.

5.2. Comparison and Analysis of Convergence Performance

5.2.1. Transverse Comparison and Analysis. Compared with other traditional algorithms such as PSO and GA, the FoSSA has obvious improvements in the convergence speed and optimization accuracy. In this study, three thirty-dimensional test functions are used to compare the convergence speed and optimization ability of FoSSA, GA, PSO, and GWO algorithms. The expressions of the test functions are shown in Table 4.

To make the results more convincing, each test function is tested 30 times independently. The population is set to 100 and the maximum number of iterations is 1000. For FoSSA,

TABLE 6: Comparison of optimization results.

Function	Optimizer	Optimal value	Worst value	Mean
$f_4(x)$	SSA	−10139.98	−6154.46	−8145.43
	FoSSA	−36168.05	−28325.41	−34049.27
	LevySSA	−30725.72	−26342.56	−28072.35
	tSSA	−29325.62	−23210.74	−28705.26
	RandSSA	−23257.82	−19072.23	−21072.43
$f_5(x)$	SSA	−15642.13	−13425.62	−14568.18
	FoSSA	−35653.03	−31971.04	−34895.24
	LevySSA	−16725.12	−14236.42	−15742.84
	tSSA	−16732.45	−14584.72	−15643.72
	RandSSA	−17325.42	−14346.28	−16435.65
$f_6(x)$	SSA	−3.2432	−3.0898	−3.1042
	FoSSA	−3.8947	−3.8628	−3.8725
	LevySSA	−3.7243	−3.2649	−3.4634
	tSSA	−3.7254	−3.6234	−3.6927
	RandSSA	−3.3274	−3.1324	−3.2736

TABLE 7: Optimal parameter combination of different classification models.

Classifier	c	g
SSA-SVM	0.12	14.50
FoSSA-SVM	4.45	0.76
LevySSA-SVM	6.31	14.56
tSSA-SVM	1.77	8.11
RandSSA-SVM	0.14	8.34

the accounts of producers and sparrows aware of danger are set as 20% and 10% of the whole population and v_s is 0.8. Crossover probability c is 0.9, mutation probability μ is 0.03 in GA, $\omega = 0.728$ and $c_1 = c_2 = 1.49442$ in PSO, $r_1, r_2 = \text{random}(0, 1)$, and $a \in (0, 2)$ with a linear decrease
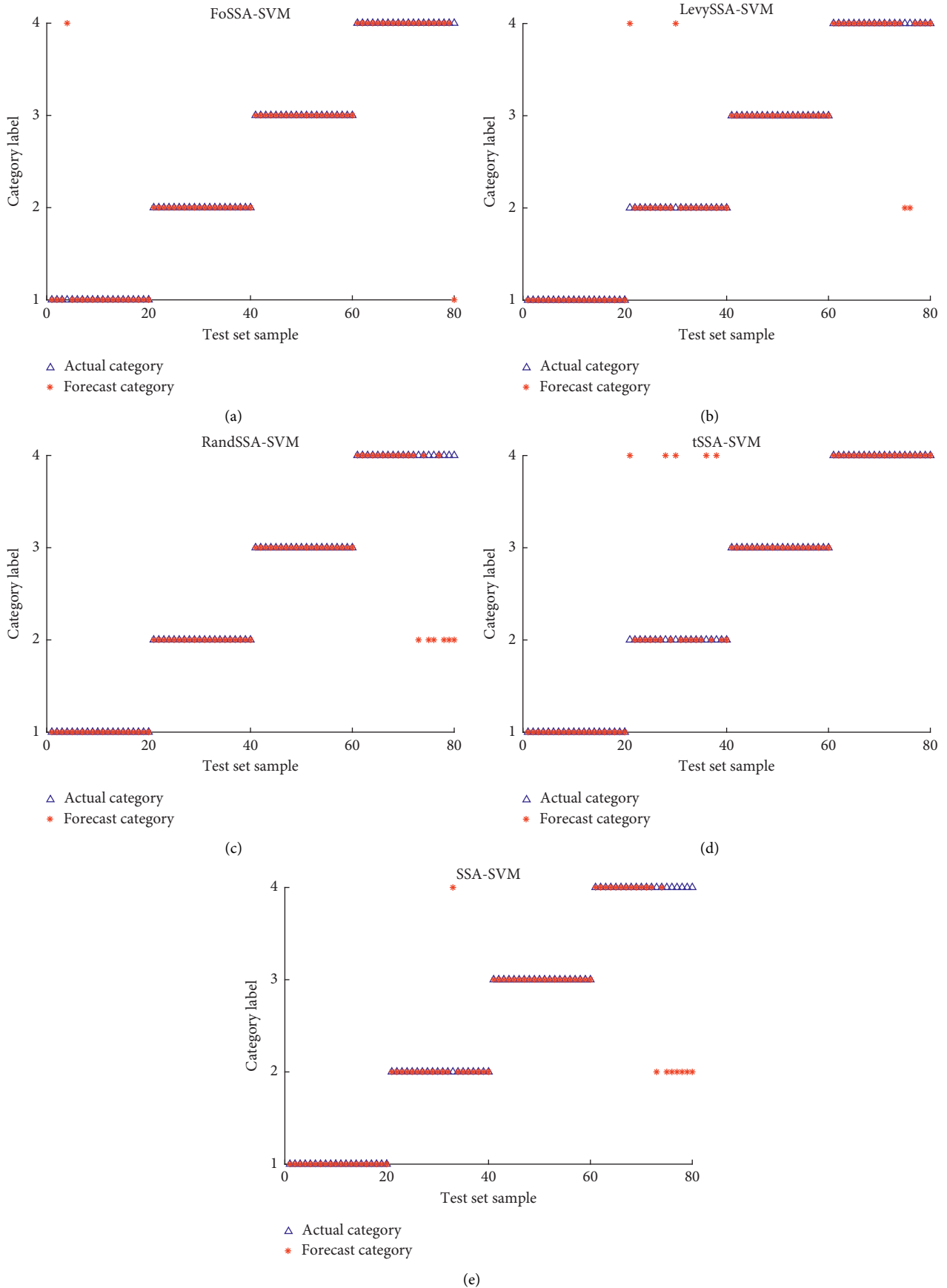
Figure 6: Classification results of different classification models. (a) FoSSA-SVM. (b) LevySSA-SVM. (c) RandSSA-SVM. (d) tSSA-SVM. (e) SSA-SVM.

in GWO. The optimization results of the four algorithms on the test function are shown in Figure 4.

The minimum values of the test functions are all zero. The results show that the PSO method has a fast convergence speed, but the convergence accuracy is very low. The GA and GWO are prone to local optimality. The FoSSA achieves the best convergence speed and convergence accuracy at the same time.

5.2.2. Longitudinal Comparison and Analysis. SSA, LevySSA [16], RandSSA [17], and tSSA [18] are studied separately to compare with our FoSSA to demonstrate the optimization effect and convergence ability. For SSA, the accounts of producers and sparrows aware of danger are set as 20% and 10% of the whole population and the safety threshold value V_s is 0.8. For FoSSA, additional parameters of the step-size factor τ is 0.2 and the maximum attractiveness ρ_0 is 2. The parameter β in LevySSA is 1.5. The parameter $r(t)$ in RandSSA is set as 0 or 1 randomly. The parameter p in t-SSA is 0.5.

In this study, a six-dimensional single-peak function is employed to test the optimization ability of the function, and two thirty-dimensional multimodal functions are employed to test the ability to escape from the local optimum. The three test functions are shown in Table 5.

Figure 5 shows the comparative analysis effect of five optimization algorithms on the minimum optimization of the test function. It is obvious that the FoSSA achieves the fastest convergence speed for the single-peak function of $f_6(x)$ and the strongest ability to transfer from local optimum for the multipeak functions of $f_4(x)$ and $f_5(x)$.

Each method optimizes 100 times for test functions, and the mean value, optimal value, and worst value are recorded, as shown in Table 6.

After comparing the FoSSA with the other four optimizers, it is found that the FoSSA obtains the strongest performance to get over the local optimum for the multipeak function and the fastest convergence speed of the single-peak function. The optimal value, worst value, and mean value of the FoSSA are the smallest, which means the best performance.

5.3. PD Pattern Recognition Results and Analysis. Based on the FoSSA algorithm, in order to find the optimal kernel function parameters g and the penalty factor c, the corresponding combination which achieves the minimum classification error rate after 30 iterations is chosen as the optimal parameter combination of g and c. The optimal parameter combination of different optimization algorithms is shown in Table 7.

The SSA-SVM, LevySSA-SVM, tSSA-SVM, RandSSA-SVM, and FoSSA-SVM classification models described in this paper are employed to recognize cable defects patterns with their optimal parameter combination. The results are shown in Figure 6 and Table 8.

Compared with the SSA-SVM, FoSSA-SVM improves the classification accuracy by 7.5%, and with the other

TABLE 8: Comparison of accuracy and times consumption of different models.

Classifier	Accuracy	t (ms)
SSA-SVM	90%	512
FoSSA-SVM	97.5%	362
LevySSA-SVM	95%	494
tSSA-SVM	93.75%	394
RandSSA-SVM	92.5%	402

TABLE 9: Prediction accuracy of different models.

Classifier	FoSSA-SVM	PSO-SVM	GA-SVM
Error number	2	7	9
Accuracy	97.5%	91.25%	88.75%

classification models, it improves the accuracy by 2%–5%. In terms of time, the FoSSA-SVM requires the shortest iteration time, which is 32–150 ms shorter than that of other algorithms. In short, the FoSSA-SVM achieves a faster optimization speed and the highest recognition accuracy at the same time.

In order to further verify the prediction accuracy, the FoSSA-SVM model is compared with PSO-SVM and GA-SVM models. 80 samples are used for testing, and each model is employed to predict 30 times; the optimal prediction result is taken into account. The final results are shown in Table 9, in which it is obvious that the FoSSA-SVM model achieves a predictive accuracy of 97.5%, which is better than that of the other two algorithms.

The results of the experiment demonstrate that the proposed FoSSA-SVM model improves the prediction accuracy significantly in cable PD pattern recognition and it achieves obvious advantages in potential applications.

6. Conclusion

FoSSA is proposed in this paper to optimize the kernel function parameters and penalty factors of SVM for PD pattern recognition of cables. A novel Circle-Gauss hybrid mapping model used in the initialization stage of SSA improved the diversity of the sparrow population. Dynamic step-size and firefly disturbance strategy help SSA out of local optimum and then improve the convergence accuracy. Compared with SVM optimized by SSA, the classification accuracy is increased by 7.5% and the time consumption is shortened by 150 ms. The introduction of firefly perturbation and dynamic step strategy enhances the global convergence ability of SSA.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the present study.

Acknowledgments

References

[1] Y. Sekiguchi, "History and recent trend of technologies on insulation materials of transmission cable," *IEEJ Transactions on Fundamentals and Materials*, vol. 139, no. 9, pp. 400–405, 2019.

[2] X. Gu, S. He, Y. Xu, Y. Yan, S. Hou, and M. Fu, "Partial discharge detection on 320 kV VSC-HVDC XLPE cable with artificial defects under DC voltage," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 25, no. 3, pp. 939–946, 2018.

[3] C. Ding, T.-Y. Bao, and H.-L. Huang, "Quantum-inspired support vector machine," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[4] K. Chang, B. K. Boyanapalli, and N. Wu, "Application of fuzzy entropy to improve feature selection for defect recognition using support vector machine in high voltage cable joints," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 27, no. 6, pp. 2147–2155, 2020.

[5] B. M. T. Fouda, B. Yang, D. Han, and B. An, "Pattern recognition of optical fiber vibration signal of the submarine cable for its safety," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6510–6519, 2021.

[6] N. Morette, T. Ditchi, and Y. Oussar, "Feature extraction and ageing state recognition using partial discharges in cables under HVDC," *Electric Power Systems Research*, vol. 178, 2020.

[7] X. Peng, J. Wen, Z. Li et al., "Rough set theory applied to pattern recognition of Partial Discharge in noise affected cable data," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 1, pp. 147–156, 2017.

[8] J. Jineeth, R. Mallepally, and T. K. Sindhu, "Classification of partial discharge sources in XLPE cables by artificial neural networks and support vector machine," in *Proceedings of the 2018 IEEE Electrical Insulation Conference (EIC)*, pp. 407–411, San Antonio, TX, USA, June 2018.

[9] F. Barani, A. Savadi, and H. S. Yazdi, "Convergence behavior of diffusion stochastic gradient descent algorithm," *Signal Processing*, vol. 183, 2021.

[10] S. Rathod, A. Ghosh, and R. Kulkarni, "Fast and accurate autofocusing algorithm in digital holography based on particle swarm optimization," *Optik*, vol. 247, 2021.

[11] X. L. Liang, R. Li, Z. H. Li, and M. Lei, "An improved ant colony optimization algorithm with negative feedback," *Journal of Physics: Conference Series*, vol. 2033, no. 1, 2021.

[12] M. Yesilbudak, "Parameter extraction of photovoltaic cells and modules using grey wolf optimizer with dimension learning-based hunting search strategy," *Energies*, vol. 14, no. 18, 2021.

[13] S. Chiang, A. K. Sangaiah, Y. Chen, and Y. Liu, "A novel artificial bee colony optimization algorithm with SVM for bio-inspired software-defined networking," *International Journal of Parallel Programming*, vol. 48, no. 2, pp. 310–328, 2018.

[14] N. Ni and Y. Zhu, "Self-adaptive bacterial foraging algorithm based on estimation of distribution," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 3, pp. 5595–5607, 2021.

[15] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow search algorithm," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 22–34, 2020.

[16] M. P. Rajakumar, J. Ramya, and B. U. Maheswari, "Health monitoring and fault prediction using a lightweight deep convolutional neural network optimized by Levy flight optimization algorithm," *Neural Computing & Applications*, vol. 33, no. 19, Article ID 12534, 2021.

[17] F. Pandey, P. Dash, and D. Sinha, "Attack-resistant and efficient cancelable codeword generation using random walk-based methods," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2025–2043, 2021.

[18] B. V. Popovic and A. I. Genç, "On extremes of two-dimensional Student-t distribution of the Marshall–Olkin type," *Mediterranean Journal of Mathematics*, vol. 15, no. 4, 2018.

Research Article

Design and Implementation of IoT Data-Driven Intelligent Law Classroom Teaching System

Gaoyan Shi (ID)

Shijiazhuang Vocational and Technical College, Shijiazhuang 050081, China

Correspondence should be addressed to Gaoyan Shi; shigy@sjzpt.edu.cn

In this paper, we conduct in-depth research and analysis by building an IoT data-driven intelligent law classroom teaching system and implementing it in the actual teaching process. Firstly, the application requirements and main functions of the classroom interactive system are analysed and studied in depth; the overall design of the classroom interactive system based on wireless communication is carried out; and the classroom interactive system consisting of teacher's receiver, wireless receiver, student's handheld terminal, teacher's receiver upper computer software, student's handheld terminal upper computer software, and data management website is developed. Platform and communication part: middleware comprises multiple modules such as real-time memory event database, task management system, and event management system. The system uses USB communication, serial time-sharing communication, Zigbee wireless communication, WebSocket, and other technologies to realize data communication between several modules, and some key technologies are studied and improved in detail. Finally, the system is deployed to the actual experimental application scenario, and the software and hardware construction of the IoT innovation lab is completed. The lab equipment is officially put into use, the information platform is also in the trial run stage, and the system is tested in time during the operation stage to provide the basis for the improvement of the system functions. Use data to conduct an in-depth analysis of learning situations and learning process analysis, optimize the reasonable construction of a personalized learning environment, serve personalized learning, and achieve the most optimal learning effect. The secondary development of the Turing robot provides a personalized auxiliary Q&A system for teaching, allowing artificial intelligence-assisted teaching to be integrated into the curriculum. With the help of teaching law courses, the empirical study demonstrates that the personalized learning environment has significant advantages in improving teaching effectiveness.

1. Introduction

The rapid development of information technology has accelerated the deep integration of information technology and education teaching, information technology platforms, and artificial intelligence-assisted teaching; other information technology resources have emerged; and a variety of new learning methods have emerged to show that education is making headway towards the information technology process. However, due to the complexity of information tools, teaching reform often tends to go astray, and very often information technology tools become just a landmark application to highlight new learning styles, staying only for information technology competitions, or open or observed classes, and not falling into the classroom reality for students

[1]. The main research purpose of this paper is to select, integrate, and develop informatization resources under a data-driven perspective and build a personalized informatization learning environment that meets the school, teaching, and learning conditions. Network nodes are required to have the self-adjustment ability, self-adaptation ability, and robustness and complete tasks such as network initialization and fault self-repair through topology control mechanisms, network protocols, and cooperation to maintain the normal operation of the network. Traditional school education cannot teach each student according to his or her ability, but the information age has brought new education opportunities [2]. With artificial intelligence and big data, the whole process of student learning can be monitored and tracked in real time, and the problems and

deficiencies of students can be analysed so that teachers can tailor their learning plans to the specific situation of students. The emphasis on process-oriented data on student learning guides students to adjust their learning styles promptly, making large-scale personalized learning possible. While offline traditional education allows teachers to impart knowledge and solidify the foundation, online education breaks the time and space constraints from which students can access personalized resources and learning styles.

Changing the traditional education model through new educational technology equipment has received increased attention from schools [3]. At present, the information management system for teachers and students has been widely used in colleges and universities, realizing the functions of integrated management of teacher and student information, online course selection, online grade entry and query, and integrated query of school information. In particular, the online grade entry and query function of students is especially widely used among teachers and students. However, the current teaching situation in domestic universities is still the traditional irrigation method interspersed with multimedia teaching methods such as PPT. Teachers still count students' knowledge proficiency through traditional after-class assignments and randomly checking students' answers to questions in class [4]. At the same time, according to the student's learning data feedback, the learning environment resources are adjusted in real time, and finally, a dynamic and personalized learning environment that conforms to the academic conditions is constructed. Once students do not keep up with the teacher's teaching ideas, it is difficult to listen to the later knowledge. As for the student's attendance, teachers usually still confirm it through the roll call.

About the methodological aspects of rule of law education, rule of law education is most often based on problem-based and case-based approaches [5]. The problem-based approach requires students to identify, construct, and analyse problems, and students, therefore, spend more time on learning activities and perform better than those in the case-based approach. Based on the needs of laboratory management and practical training teaching, it is necessary to research and explore laboratory intelligence and informatization for the management of university laboratories. If the requirements are met, open the serial port communication switch for communication, and close the serial port switch after the communication is completed. If it does not meet the requirements, wait for a period, continue to query until there is no other coordinator communication, and then turn on the serial port switch. In this paper, based on the existing perfect facilities, the structure and characteristics of IoT technology in smart home and smart building are applied to the laboratory management system, exploring the model of intelligent laboratory, constructing IoT innovation training room, and developing smart campus class terminal management system, which integrates the use of IoT, mobile Internet, big data, intelligent perception, and other emerging information technologies, on the basis of traditional campus, to provide timely feedback to parents and teachers on students' safe attendance and abnormal attendance; to

organically combine all aspects of home-school interaction, campus information, announcement release, class culture, teacher-student interaction, class walking teaching and teaching management; to build an information management platform for campus and class management and student growth; enrich the overall school information technology environment and information dissemination channels; efficiently and conveniently deliver various knowledge information and notice information; provide timely feedback on classroom information; realize transparent management of classes; and provide strong technical support for the digitization and information of education and teaching in our school.

2. Status of Research

The spreading factor and the communication rate are found to have a significant impact on the coverage capability of the LoRa network through field tests. The limitations of the Lora WAN protocol are proposed through the analysis of Lora WAN protocol, the size of the Lora WAN network is much smaller than the theoretical value under the limitation of duty cycle transmission, and the adoption of transmission acknowledgment answering mechanism for reliable transmission also makes the network capacity smaller in the same way [6]. Through different IoT application cases, the advantages and disadvantages of Lora WAN in different scenarios are analysed and its applicability scope is illustrated. The authors argue that the random-access method of Aloha for nodes in Lora WAN networks is not the optimal solution and a more complex TDMA access method can be considered. By modelling and simulating a large Lora WAN network using the NS-3 network simulator, the simulation results show that the limited downlink communication in the Lora WAN network reduces the packet delivery rate of the uplink communication that needs to be acknowledged and adding gateways can improve but not eliminate this effect [7]. A comprehensive analysis of the features and limitations of Lora WAN gives the current problems of the Lora WAN protocol, that is, high data loss rate at high data concurrency and long retransmission time for the timely arrival of data [8]. The physical layer transmission characteristics of the Lora WAN network are analysed, the packet loss rate and conflict probability of different data frames at the gateway are derived by building a signal reception model, and the variation of packet loss rate and conflict probability with different spreading factors and sending cycles are illustrated by simulation results [9].

With the development of communication technology, classroom interactive systems have gradually entered university campuses and are playing an increasingly important role in student learning. Multimedia teaching has also become a standard part of daily university teaching activities [10]. Many university research project groups and some social research institutions at home and abroad have also conducted a series of research on education informatization [11]. There are currently two ideas for developing classroom interactive systems. One idea is to upgrade the hardware facilities completely and transform them into intelligent

classrooms [12]. However, these devices require a large-scale and comprehensive upgrade of the classroom hardware; the amount of engineering, time costs, and transformation costs are relatively large. Moreover, if a certain aspect needs to be upgraded at a later stage, it will be more difficult [13]. The data management web terminal is mainly used to bind the user's school number and the MAC address of the student's handheld terminal. The teacher downloads the student list and the corresponding MAC address list of the class to be taught, stores and analyses the students' classroom answer and question data collected by the teacher, and visualizes these data. For the student user, it can display the students' usual classroom performance in real time and show it in a chart; for the teacher user, it can display his classroom performance in real time [14]. For teachers, they can display the class performance of students in their classes in real time and display it in charts [15].

With the continuous development and updating of computer technology, big data, cloud computing, and artificial intelligence will be widely used in all areas of social life. For example, big data can be used to predict the occurrence of crimes; companies can analyse the sales of products based on customer reviews on shopping websites; the National Centre for Disease Control and Prevention can analyse the spread of diseases based on Internet users' search records; and so on [16]. For the large amount of data generated, the integration and analysis of the data allow people to discover new knowledge and create new value. As the volume of IoT business increases, the amount of data storage and computation will bring the need for "computing" power to the cloud. Cloud computing will provide the appropriate services in infrastructure, platforms, and software. Artificial intelligence is a branch of computer science and new technical science that combines the application of computers and other disciplines, which focuses on the study of theories, methods, and technologies to simulate, extend, and expand human intelligence [9].

3. IoT Data-Driven Intelligent Law Classroom Teaching System Design

3.1. IoT Smart Classroom Teaching System Design. There is no doubt that the communication among many nodes in IoT involves addressing the problem, and IPv4 addresses are getting scarce. Its address space can hardly meet the demand for address space in today's IoT. Moreover, IPv4 has a shortcoming for the mobility of the Internet, and the MIPv4 mechanism proposed in the IETF to support the mobility of nodes will lead to rapid consumption of network resources for many nodes, resulting in network paralysis [17]. IPv6 technology not only has a huge address space but also adopts a stateless address allocation scheme to solve the problem of allocation efficiency of massive addresses. For node mobility, IPv6 proposes the concept of IP address binding buffer and introduces a method to detect node movement. At the same time, IPv6 design considers the quality of service, such as defining the traffic class field and flow label field in the packet structure and using the flow label carried by the packet to do a detailed division of services. It is a protocol designed for

communication between remote sensors and control equipment that have limited computing and storage resources and work on low-bandwidth, unreliable networks. It has the characteristics of simplicity, strong scalability, and low traffic.

The IoT application layer is the ultimate purpose layer, which is the combination of IoT and industry needs, by docking IoT technology with industry technology and using various solutions to provide a good IoT business experience for a wide range of users to achieve a wide range of intelligent applications, so that people can feel the great impact that IoT brings to real life. The main function of the IoT application layer is to process the massive information coming from the network layer and use this information to provide relevant services to users. The key issue of the application layer is how to achieve information sharing and security. Among them, the rational use, as well as efficient processing of relevant information, is an urgent IoT problem, and to solve this technical challenge, the IoT application layer needs to utilize technologies such as middleware, Machine to Machine (M2M), and edge cloud computing. In building the information network of IoT, middleware is mainly applied to distributed systems to connect various technologies and share resources. Middleware can be divided into two parts: the platform and the communication part. The middleware consists of various modules such as a real-time memory event database, task management system, and event management system, which can be implemented to connect two independent applications, as shown in Figure 1.

It is a protocol designed for communication between remote sensors and control equipment that have limited computing and storage resources and work on low-bandwidth, unreliable networks. It has the characteristics of simplicity, strong scalability, and low traffic. A wireless sensor network consists of tiny nodes deployed in the monitoring area with data processing units, miniature sensors, and communication modules through a self-organizing wireless self-organizing network system. With the help of the nodes' built-in sensors measuring the signals of the objects around their location, the observer can obtain information about the sensed objects in the network coverage area. The three elements of wireless sensor networks are sensor, sensing object, and observer [18]. The characteristics of wireless sensor networks are mainly reflected in the size of scale and the purpose of application. In the application of a wireless sensor network, the location of sensor nodes cannot be set precisely in advance. Therefore, the network nodes are required to have self-regulation ability, self-adaptive ability, and robustness to maintain the normal operation of the network through topology control mechanism and network protocols and mutual collaboration to complete such work as network initialization and fault self-healing. The basic structure of wireless sensor networks can be divided into node types, node structures, and function-based models of wireless sensor network structures. Its technical research is to be mainly focused on medium access control protocols, network topology control, routing protocols, node positioning, clock synchronization, data management, data fusion, embedded operating systems, and network security.
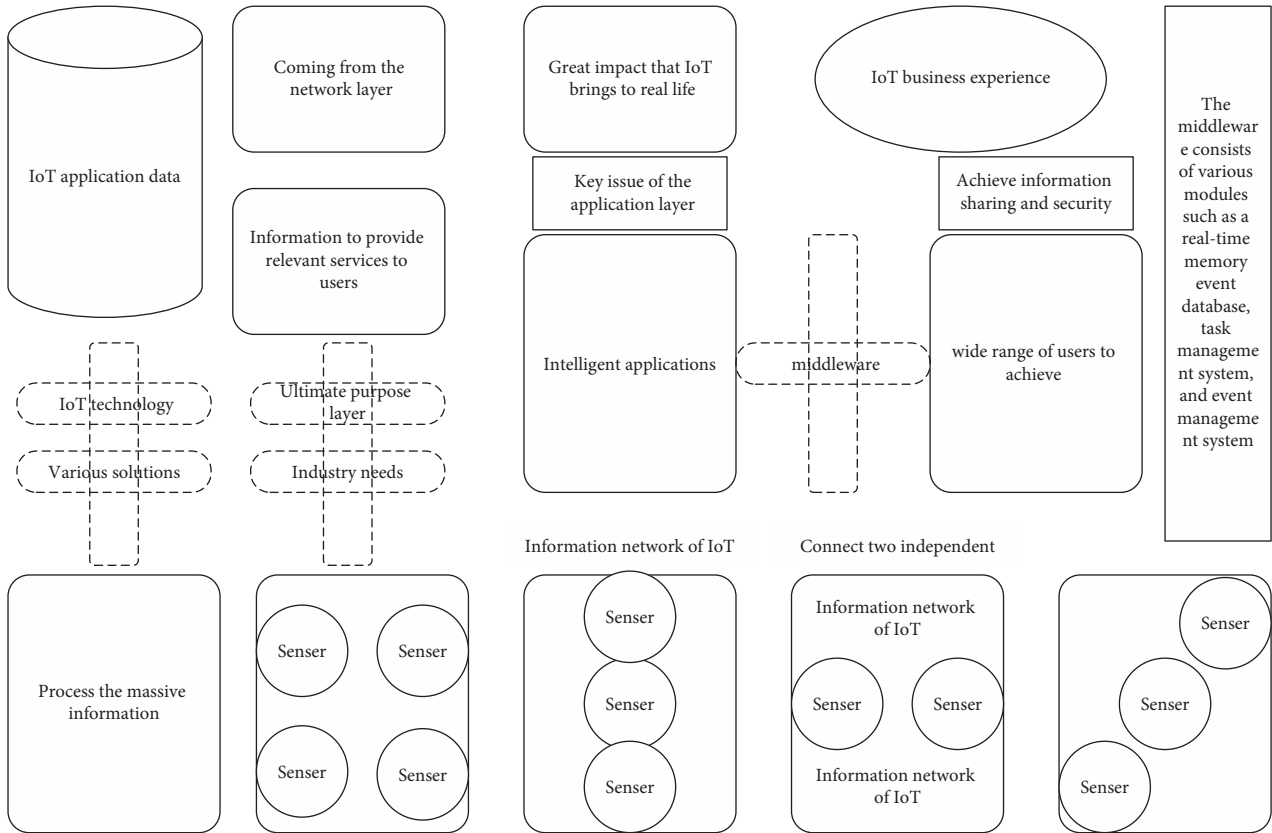
Figure 1: IoT framework design.

Based on the in-depth analysis of constructivist theory, we propose the theory of realm pulse learning and build a personalized learning environment with an information platform and AI teaching assistance. It is a new paradigm to guide the transformation of learning, which allows the interaction between the many learning elements of the learner and the external elements of the teacher's guidance to form a self-renewal force. In many cases, informatization methods have only become an iconic application for demonstrating new ways of learning. They only serve for informatization competitions or open classes and observation classes and have not been implemented in the classroom to serve students. The AI-assisted teaching of the information platform can undoubtedly provide a better learning environment for the student-cantered teaching model while adjusting the learning environment resources in real time based on the feedback of students' learning data, finally building a dynamic and personalized learning environment in line with the learning situation.

Then, to make data-driven decision-making applied to teaching in addition to the basic elements, some conditions are needed. One is based on the concept of big data teaching, education data as a practical tool, and means to promote the improvement and perfection of teaching continuously. Second, for the correct understanding of the application of data in teaching, the big data should have a cautious attitude; there may be some personal privacy, data security, and other issues; and we should adhere to the principle of science and objectivity, combined with the actual situation and flexible

choice. The third is the construction of an effective platform for data collection and application, to be able to use the platform to carry out teaching activities inside and outside the classroom, collect learning data, carry out learning tracking, and make data decisions truly applied to teaching, as shown in Figure 2.

To ensure that more student handhelds can be connected in a wireless-based classroom interactive system, we need a teacher receiver and 4 Zigbee coordinators to transmit information to each other. Reasonably use educational data to customize a personalized learning environment to meet the learning needs of students at different levels so that artificial intelligence and big data can truly promote more personalization, thereby increasing the attractiveness of the classroom. The teacher receiver uses one serial port to connect to 4 wireless receivers, respectively. At the same time, there are 4 LCD lights connected to each hardware channel to show which coordinator of the wireless receiver the teacher receiver is communicating with. There is also a serial communication switch connected to each serial path, which can only be communicated when the switch is closed. When the teacher needs to send a message to the 4 serial ports with the receiver, the 4 serial communication switches can be turned on and off in sequence to ensure that only one serial communication switch is on for each communication. When a coordinator needs to send a message to the teacher receiver, it also needs to make sure no other coordinator is communicating first. If the requirements are met, then turn on the serial communication switch to communicate and
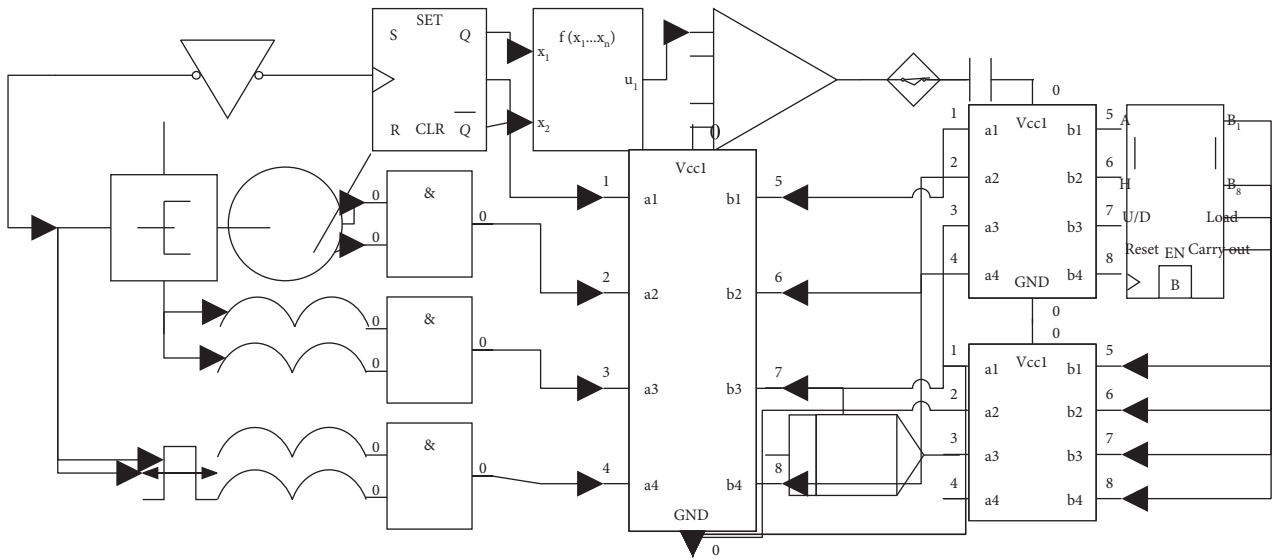
FIGURE 2: Time-sharing communication serial port principle.

turn off that serial switch when the communication is complete. If the requirement is not met, then wait for some time and continue to query until no other coordinator is communicating before turning on the serial switch.

Reasonably use educational data to customize a personalized learning environment to meet the learning needs of students at different levels so that artificial intelligence and big data can truly promote more personalization, thereby increasing the attractiveness of the classroom. A typical Lora WAN network architecture consists of four parts: nodes, gateways, web servers, and application servers. The nodes are mainly responsible for the collection and reporting of sensing data and the reception and execution of command data on the server side. The nodes adopt the ALOHA access method to randomly access the network and send data on demand. The data reported by a node can be received by multiple gateways, thus forming a star network topology with the gateways, which can ensure the reliability of the connection and realize the access of a larger number of devices in the network. The gateway is the relay device between the node and the network server in the Lora WAN network as the central node of the star network, forwarding the LoRa RF data from/to the node and TCP/P network data for processing. The network server is an important part of the Lora WAN network for data processing and device access, responsible for LoRa node data parsing and encapsulation, encryption, decryption, LoRa node entry authentication, Adaptive Date Rate (ADR) regulation, and other functions. The application server is built and developed by the user side and has different application functions for different IoT applications. Building a database to process and store the application data, providing an application logic interface, and using a visual interface such as WEB or APP to connect the user side with the application server mean building a LoRa IoT system from the sensing layer to the application layer.

3.2. Data-Driven Intelligent Law Classroom Analysis.
Depending on the starting point for considering the problem, research scholars have proposed model-driven, goal-driven, product-driven, business-driven, technology-driven, decision-driven, task-driven, and other forms of drivers. Data analysis emanating from the model side is called model-driven, and the mode of operation with the technology side as the trigger is called technology-driven. With the advent of the intelligence and data era, knowledge acquisition is no longer entirely dependent on expert experience. The use of some intelligent methods can start from the data, mining to get the valuable information behind the data, which is data-driven and one of the main reasons why data-driven methods have developed more rapidly than other methods and attracted widespread attention from scholars [19]. From the perspective of the generation environment, the concept of data-driven is a product of the era of big data, inspired by data-intensive scientific discoveries. Data has become the starting point and perspective for people to understand, analyse and solve problems, resulting in the idea of using a data mindset to solve problems. Furthermore, the scientific preprocessing of data, the constant statistics of data, the real calculation of data, data modelling, machine learning, data visualization, data management, and other technologies provide the technical support for data-driven, in addition to these data science methods, data mining, neural networks, artificial intelligence, metadata, and all other technologies and methods of data-based design are the technical guarantee of data-driven. Reasonably use educational data to customize a personalized learning environment to meet the learning needs of students at different levels so that artificial intelligence and big data can truly promote more personalization, thereby increasing the attractiveness of the classroom:

$$T_p = \frac{\sqrt{x^2 + y^2 + z^2}}{C},$$

$$T_s = \frac{L}{V} - \frac{\sqrt{x^2 - y^2 - z^2}}{C}. \tag{1}$$

The essence and value of big data are to eliminate uncertainty, and the value of data drive lies in the kinetic output of the drive and the means to realize the value of big data: data analytic thinking, data-driven applications in different scenarios, the formation of data-driven decision-making, data-driven design, data-driven procedures, data-driven products, data-driven innovation, and other results output. Among them, the so-called data-driven decision-making is to help patronize through data, including product improvement, operation optimization, marketing analysis, and business decision-making. With data support, we can better determine which channels convert better and which feature styles are more popular with users and support the decision through data. Data-driven product intelligence is the ability to empower products to learn through data models. The so-called intelligence can be understood as a model with a certain database, setting an algorithm model on it; then, the results of the data obtained back to the product are intelligent. In this way, the product has a certain learning ability and can constantly update and iterate itself, for example, the literature recommendation system, by analysing the user's historical behaviour records, discovering the user's interest points, building a targeted user model, and finally realizing personalized recommendations, as shown in Figure 3.

Scientific discovery has shifted from the traditional hypothesis-driven to a scientific approach based on scientific data for exploration. For some scientific problems that are difficult for humans to understand, researchers are gradually unravelling the mystery of these scientific puzzles through data monitoring and analysis, using data as a tool and object of research and designing and conducting research from data. Data are no longer just the result of scientific research but have become the living foundation of scientific research. People are concerned with not only modelling, describing, organizing, preserving, accessing, analysing, reusing, and building scientific data infrastructure but also how to use the ubiquitous network and its inherent interactivity and openness to construct an open and collaborative discovery and creation model with the help of massive data, thus giving birth to data-intensive scientific discovery. There is a lack of effective and detailed information interaction between students and teachers. Teachers cannot specifically and comprehensively understand students' knowledge proficiency. They can only arrange teaching content and teaching progress based on their own teaching experience and random checks in the classroom. The development of a data-intensive scientific research paradigm offers new opportunities for collaborative efforts between academia, industry, and government. This research paradigm relies on sensing tools to acquire data or computer

simulations to generate it, process it through software for cleaning, store it in computers, and analyse it using data analysis software. Data from scientific research can be grouped into four categories: data generated by experimental instruments, data generated by computer simulations, data collected by observation, and data generated on the Web. Data generated on the Web include historical behavioural data of Internet users and data on Web transactions:

$$a(h_m) = (1.25gf + 0.7)h_m - (1.25gf - 0.7),$$

$$P_{\text{coll},sf} = 1 + e^{2G_{sf}}. \tag{2}$$

Data visualization is an interactive visual representation of data using the intelligence of the human brain and the perceptive ability of the human eye to convert data that is difficult to display directly into graphics, symbols, and so on to convey useful information efficiently. Through this visual representation of data, the information hidden in the data, including various attributes, is presented in some summary form. From the point of view of aesthetic requirements and functional requirements, data visualization requires the same for both and guarantees a balance between design and functionality by directly expressing special aspects and special features, and thus gaining insight into very sparse but particularly complex data collections.

The conflict model of the Lora WAN protocol shows that the probability of conflict during transmission of data packets sent by nodes using ALOHA access is closely related to the combination of bandwidth, coding rate, and the proportion of nodes using the same spreading factor. When the Lora WAN protocol is simulated in the LoRa network created in this paper, the corresponding proportion of nodes using the same spreading factor, bandwidth, coding rate, and other parameters are substituted into the formula, and the probability of conflicts occurring in the data transmission process of nodes using different spreading factors is close to 1. At this time, the conflicts in the network are more frequent, the data loss caused by communication conflicts is more serious, and the reliability of network communication is poor at this time:

$$G_{sf} = \lambda^2 \partial_{sf} T_{sf}. \tag{3}$$

As the information transfer channel in the whole IoT architecture, the network layer is responsible for transferring the information obtained from the sensing layer and submitting it to the application layer for processing. The application layer is the interface between the IoT system and the user, which mainly solves the problems related to information interaction, human-machine processing, and so on. It combines the needs of various IoT industries to complete specific business functions and realize the intelligent application of the whole IoT industry. The main tasks of the application layer include data analysis and processing, data storage and mining, and communication handover between people and things, and things and things [20]. To adapt to the ubiquitous situation of devices and at the same time play the powerful computing ability, storage ability, and network communication ability of cloud
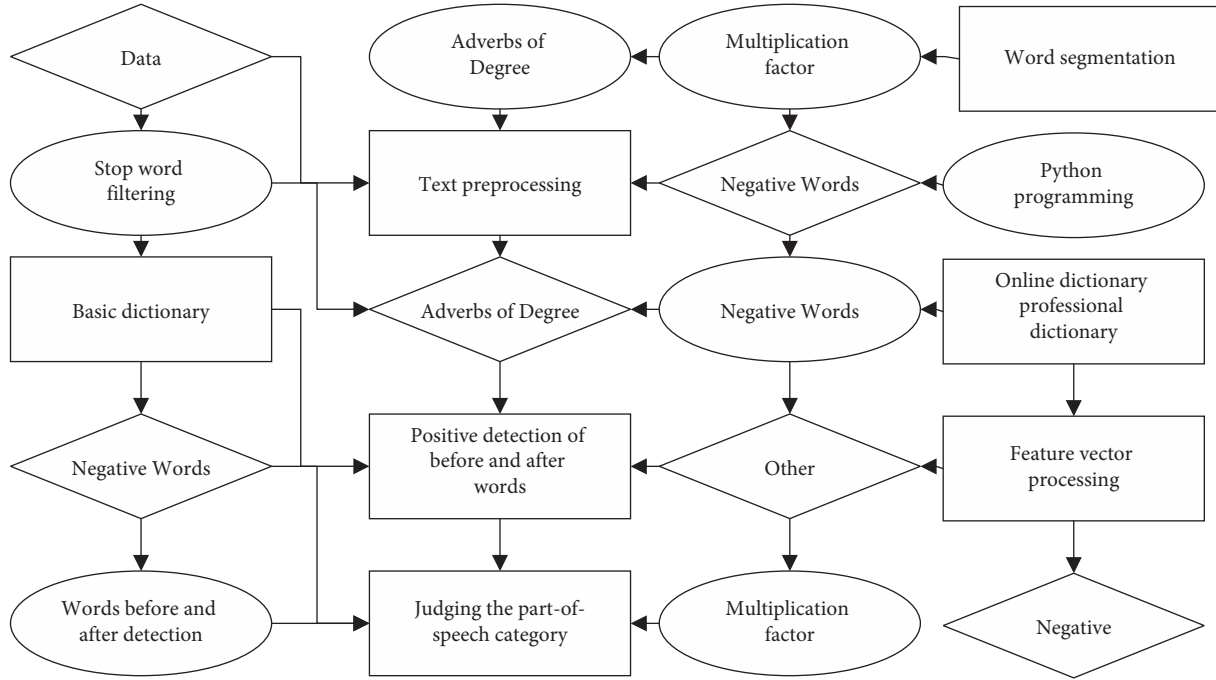
FIGURE 3: Data-driven intelligent law classroom architecture.

computing platform, the cloud platform is used as the application layer in the IoT architecture to build the IoT cloud platform architecture; establish the unified management of people and things in the IoT application; establish and maintain the communication channel between people and things; and realize the functions of remote control, data management, business logic processing of industry applications, and so forth. The IoT system architecture based on the cloud platform is shown in Figure 4, which mainly consists of terminal layer, network access layer, and cloud platform.

For aesthetic effect, adjust the label size to be <10 so that the information about the research user can be better presented. The base feature is more important in understanding the researcher's information and is also the least variable. Set the base feature label size to 10. It is necessary to research and explore the intelligence and informatization of laboratories. This article takes the Computer Engineering Department of Fujian Institute of Information Technology as an example. Based on the existing facilities, the structure and characteristics of the Internet of Things technology in smart homes and smart buildings are applied to the laboratory management system:

$$i_t = \sigma\left(W_i \left|h_t x_t^2\right| + b_i\right). \tag{4}$$

Research preference label and research relationship label: the user portrait of the digital library knowledge discovery platform designed in this paper is essentially a labeled formal description of the user model of the platform users, and the user portrait is constructed using the user's basic attributes, research results, and search target information and displayed through the word cloud visualization. Finally, the real users of the knowledge discovery platform are selected for the method validation. The user portrait of the digital
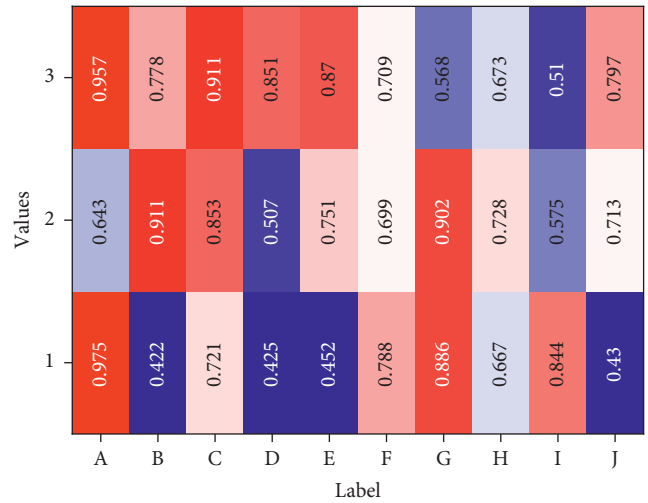
FIGURE 4: Label weights.

library knowledge discovery platform constructed in this paper can depict user information in all aspects and intuitively display user information to accurately provide personalized services and accurate recommendations and provide a real and effective reference basis for the decision-making layer.

4. Results and Analysis

4.1. IoT Smart Classroom Teaching System Performance. The key to achieving remote management from the cloud to the device side is to establish and maintain instantaneous and reliable communication between the cloud and the device side. Instant communication requires that the device

side can report data to the cloud, receive messages from the cloud in real time, and ensure that the communication has low latency. Since in the actual application environment, the network environment at the device side cannot ensure stability and reliability, so based on establishing the connection between the device side and the cloud side, a heartbeat mechanism is required to maintain the long connection from the device side to the cloud side, and at the same time, TCP-based protocols are preferred to select the communication protocols to ensure the reliability of the connection. Based on the above analysis, this paper selects the MQTT protocol to solve the problem of end-to-end instant and reliable communication during remote communication. MQTT protocol is a lightweight, agent-based publish/subscribe model messaging protocol designed for communication of remote sensors and control devices with limited computing and storage resources and working in low bandwidth and unreliable networks, with features of simplicity, scalability, and traffic saving. The MQTT protocol enables network communication between multiple clients by specifying topic subscriptions and message distribution, with the server acting as a proxy to forward messages based on the topic number. By decoupling the senders and receivers of messages in space and time, the system has good scalability and scaling capability.

The rapid development of IoT applications has led to the emergence of various types of IoT products. In various IoT products, the nodes can be either sensing devices/actuation devices with sensing data collection modules and control modules or subsystem devices connected through field buses. By accessing different functional module devices, different IoT products can be quickly constructed to achieve functional expansion of IoT applications. For the characteristics of multiple and complex product functions in IoT applications, this paper completes the description of device functions by defining attributes, services, and events for devices, referred to as the object model, to construct data models of device entities under different products, digitize the entity objects in physical space, and digitally represent them in the cloud, as shown in Table 1 for the definition of attributes and services events. By linking the Internet of Things technology with industry technologies, various solutions are used to provide users with a good Internet of Things business experience to achieve a wide range of intelligent applications, so that people can truly feel the huge impact of the Internet of Things on real life.

Different IoT products are used in different application scenarios, and the hardware architecture and software environment of the devices in the products, as well as their functions, show great differences, which determine different forms and functions of the device side, leading to the gradual "fragmentation" of IoT applications. Any user-oriented IoT product needs to break the full-stack communication protocol of "device–network platform application," but the "fragmented" IoT feature leads to great differences in the format, length, and precision of data from different devices. The communication protocols of different products are also very different, which will lead to the problem of interconnection between IoT products and become a bottleneck for

further development of IoT. In the traditional IoT application development mode, first, the device developer needs to define the communication protocol from the device side to the application side according to the system function. The application side has a strong dependency on the device side development, so the device side must be done first and then the application side, which cannot be developed synchronously. The application developer must wait for the device developer to complete the development of the device side and then carry out the development and testing of the application side, which undoubtedly increases the development time and cost and consumes huge human and material resources, as shown in Figure 5. Data-driven decision-making, data-driven design, data-driven programs, data-driven products, and data-driven innovations are formed.

In order to avoid the communication conflict caused by multiple nodes communicating with the gateway at the same time on the same channel, CSMA presentation monitoring and random delay backoff strategy are used to avoid the conflict. Presend listening is to detect the presence of packet leading code on the channel through CAD. If the CAD detects the presence of leading code transmission in the current channel, then the current channel is occupied. Otherwise, the current channel is in an idle state. If the channel is idle, then the service transmission can be carried out. Otherwise, random back-off is performed. When the channel is occupied, the timer is set to wait for a random delay; after the delay, the current channel state is detected until the channel is idle and the data is sent successfully, or the number of random evasions reaches the upper limit. If the number of random evasions reaches the upper limit, the competing channel is randomly switched, and the data is sent again.

Since the SNR values of the three groups of nodes to the gateway are similar in this test, the gateway assigns the three nodes to the same TDMA channel after admission. They are assigned communication time slots consecutively in the order of admission. The three nodes continuously occupy the time slots for periodic service reporting. Use data as research tools and objects, design, and conduct research from data. Data is no longer just the result of scientific research but becomes the living foundation of scientific research. Theoretically, there is a difference of 3 seconds between the time slots of each node for periodic reporting. However, the test data shows that there is a delay of several tens of milliseconds due to the software and hardware factors. The corresponding communication cycle has a certain delay, but in general, it is verified that the LoRa network communication with TDMA access is feasible, and different nodes will actively report at different times without conflict with each other to ensure the reliability of communication is ensured.

4.2. Results of the Data-Driven Smart Law Classroom Experiment. Finally, classroom observations revealed a strong interest in the course, most notably in the reluctance of students to leave the classroom after class and continue to

TABLE 1: Properties, events, and service definitions.

Project	Describe	Value
Attributes	Parameters describing the operating state of the equipment	13
Service	The current temperature and humidity values collected by the temperature and humidity sensing device	14
Event	The attributes of the device support read-only, write-only, read and write, and users can read and set them	54
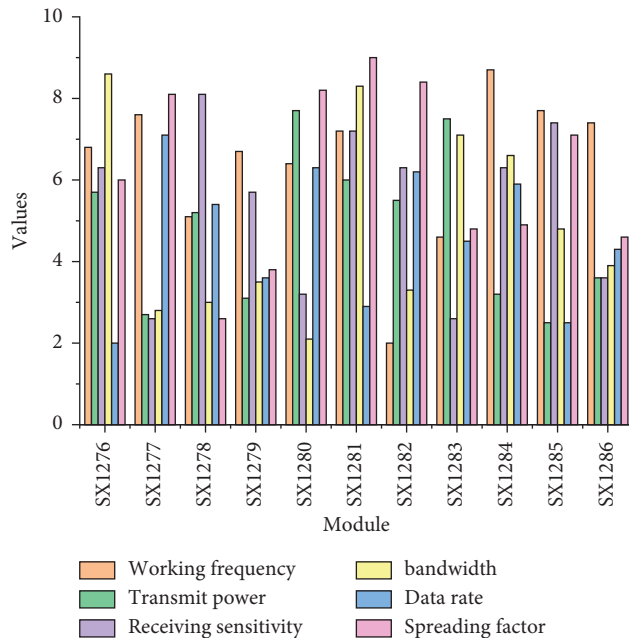
FIGURE 5: Comparison of the communication bloc.

modify the code to create their work. Several problems were also revealed: firstly, the lack of basic computer skills due to the low exposure to computers, which was a major obstacle to the completion of the first three weeks of the assignment, and secondly, the repeated shouting at the teacher during the class. In the teacher workshop after the class, we analysed two reasons behind this: students were afraid of using the software badly because they were not familiar with the use of Scratch programming, so they were afraid to try boldly when there were problems with the code; students needed the teacher to affirm the quality of their work and encourage and evaluate it before they would proceed to the next stage of the task after completing their work. Show it in some form of summary. From the perspective of aesthetic requirements and functional requirements, the requirements of data visualization are the same for both, and by directly expressing special aspects and special characteristics, we can gain insight into very sparse but particularly complex data sets to ensure balance between design and function. The data collected was divided into three parts: the first part was about mathematical achievement, evaluated with a pre- and posttest of mathematical academic achievement related to project progress; the second part was computational thinking, divided into three areas (computational concepts and computational practices were evaluated by Bibras competition questions, computational perspectives were tested by scales, and collaborative perceptions in computational perspectives were analysed separately); and the third

area was a problem-driven learning environment perceptual analysis, as shown in Figure 6.

Use the template to design a display page for the new seasonal products of Sun King Sportswear. According to the web design imitation + microinnovation concept, students can imitate the teacher's work or modify their preclass work. Moreover, according to their actual learning situation, they can choose a learning method that suits them; for example, students with weaker abilities can log on to the Super Star platform to watch tutorials and microlessons on demand and dialogue with robot assistants to help them learn, keep up with the pace of the class, and exercise their independent inquiry skills. Students upload the designed layout in real time to an online discussion group to exchange ideas with group members. This session uses online resources and a robot assistant to assist students in their task exploration, exercising their independent inquiry skills. Students discuss the layout online to improve their collaborative learning skills. The "shake it up" spot check of task completion adds interest to the classroom and instantly determines the level of proficiency in key content, recording student errors for more intuitive analysis and summary. The information resources include Super Star Online Learning Space, Turing Robot, Learning Pass "Shake One," and Super Star Interactive Group. After the teacher's lecture, students can access the online learning space to learn more or ask questions to the robot assistant according to their proficiency. In this session, students can access the Superstar platform video to watch the learning data and discussion, and the students' personalized learning mode is greatly reflected. Online group interaction makes group learning more convenient. The use of "shake" instead of roll call to check the completion of tasks increases the interest of the classroom and improves students' participation in the classroom, as shown in Figure 7.

As can be seen from Figure 7, the accuracy of the research object is 0.90, the recall is 0.84, and the F1 value is 0.87. The research object category is higher than the other categories in all indicators of discrimination because the research object generally has an established authorized name. For example, in the case of research objects in specific feature domains, the corresponding conventions can usually be found, and the model can better identify and determine them. Research questions and research methods, on the contrary, are more variable and complex in their formal composition, so the discriminative F1 values are relatively low. The change of knowledge discovery services in data-driven digital libraries has put forward higher requirements on the semantic understanding of data resource contents, especially the full text of scientific research papers. Complete specific business functions and realize intelligent applications in the entire Internet of Things industry. The main
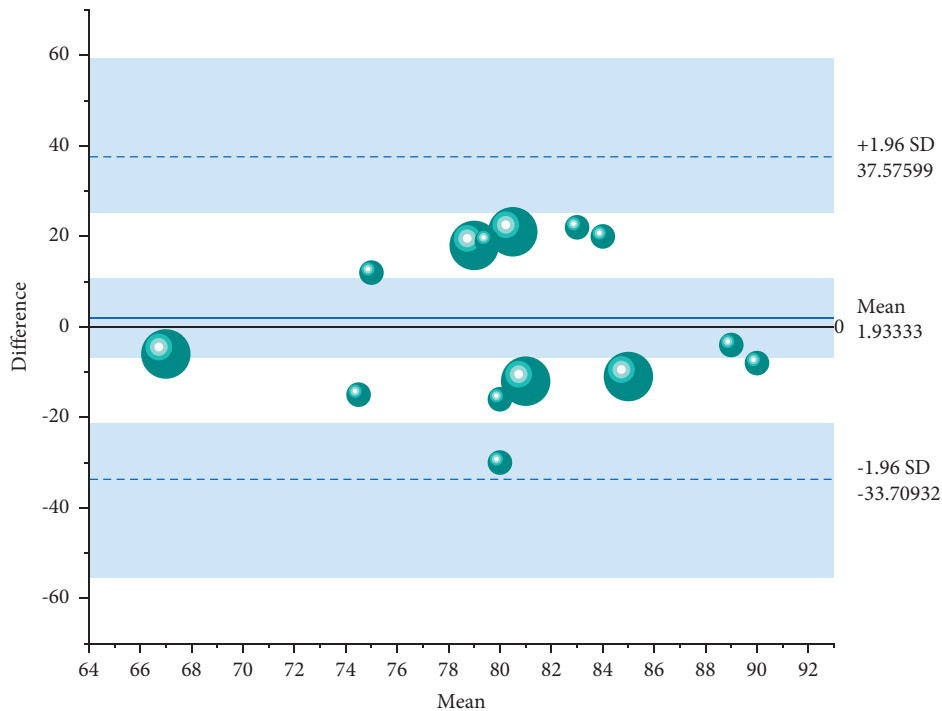
FIGURE 6: Test performance statistics.
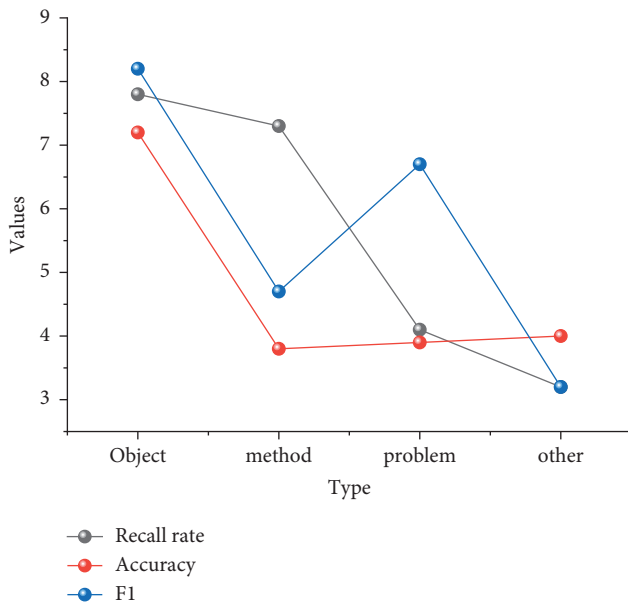
FIGURE 7: Comparison of inspection results.

tasks of the application layer include data analysis and processing, data storage and mining, communication between people and things, and things and things. To assist in achieving semantic identification, annotation, and understanding of data resource content, this section designs a keyword-based approach to identify the problems, methods, and research objects of data resource content using keywords of full-text scientific research papers and implements the approach in the field of agriculture. Experiments show that the method achieves satisfactory identification results.

This section focuses only on identifying the research object, problem, and method vocabulary, without exploring the functions assumed by the vocabulary in the text and whether the problem corresponds to the method. Therefore, future research will further explore which specific function the vocabulary assumes in the text and what the corresponding solution to the problem is. Through the experiments, the effectiveness of the deep learning model is verified, which provides an effective general idea for the research design fingerprint extraction of the full text of scientific papers in the knowledge discovery service platform of the digital library and methodological references for other types of data content extraction and identification studies.

5. Conclusion

The LoRa IoT system is designed with the help of IoT platform, the access process from the device side to the cloud side is developed, and the authentication mechanism is used to ensure the legal access of the device side. The front-end web page of the data management website design is not beautiful enough, and the functions are not perfect. The jQuery and Bootstrap frameworks used in developing the front-end of the web page are not beautiful enough, and the website currently only does some simple data processing and visualization. Now, it does not support the function of big data analysis. The MQTT protocol is used to achieve reliable remote long connection communication between the cloud and the device; the device function is abstracted through the object model, and the object model is described using the JSON data interaction format so that the relevant communication protocol can be developed to realize the data transmission between the end and the end. The wireless

communication-based classroom interaction system designed in this paper can help teachers realize the collection of feedback data, data visualization display and storage of students' roll call, multiple-choice questions of classroom quizzes, and Y/N judgment questions in classrooms. The data management website also allows students and teachers to view the actual classroom performance of students in their class or the teacher's class. Although the system can achieve the basic functions of the classroom interactive system, due to the limited personal ability, there are still numerous shortcomings, and there are some gaps between the research work of the thesis and the actual use of the needs, which need to be continuously improved at a later stage.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest regarding this article.

Acknowledgments

References

[1] M. Kwet and P. Prinsloo, "The "smart" classroom: a new frontier in the age of the smart university," *Teaching in Higher Education*, vol. 25, no. 4, pp. 510–526, 2020.

[2] Z. Engin, J. van Dijk, T. Lan et al., "Data-driven urban management: mapping the landscape," *Journal of Urban Management*, vol. 9, no. 2, pp. 140–150, 2020.

[3] V. Hillman and V. Ganesh, "Kratos: a solution for data privacy, literacy, and student agency in a data driven school ecosystem," *International Journal of Innovation in Education*, vol. 6, no. 3-4, pp. 184–201, 2020.

[4] A. Bennett, "Autonomous vehicle driving algorithms and smart mobility technologies in big data-driven transportation planning and engineering," *Contemporary Readings in Law and Social Justice*, vol. 13, no. 1, pp. 20–29, 2021.

[5] Z. Y. Dong, Y. Zhang, C. Yip, S. Swift, and K. Beswick, "Smart campus: definition, framework, technologies, and services," *IET Smart Cities*, vol. 2, no. 1, pp. 43–54, 2020.

[6] A. M. Shahat Osman and A. Elragal, "Smart cities and big data analytics: a data-driven decision-making use case," *Smart Cities*, vol. 4, no. 1, pp. 286–313, 2021.

[7] S. Y. Hong and Y. H. Hwang, "Design and implementation for iort based remote control robot using block-based programming," *Issues in Information Systems*, vol. 21, no. 4, pp. 317–330, 2020.

[8] K. Coatney and M. Poliak, "Cognitive decision-making algorithms, internet of things smart devices, and sustainable organizational performance in Industry 4.0-based manufacturing systems," *Journal of Self-Governance and Management Economics*, vol. 8, no. 4, pp. 9–18, 2020.

[9] N. Petrovic and D. Kocic, "Data-driven framework for energy-efficient smart cities," *Serbian Journal of Electrical Engineering*, vol. 17, no. 1, pp. 41–63, 2020.

[10] C. Edwards, "Real-time advanced analytics, automated production systems, and smart industrial value creation in sustainable manufacturing internet of things," *Journal of Self-Governance and Management Economics*, vol. 9, no. 2, pp. 32–41, 2021.

[11] T. Cerquitelli, D. J. Pagliari, A. Calimera et al., "Manufacturing as a data-driven practice: methodologies, technologies, and tools," *Proceedings of the IEEE*, vol. 109, no. 4, pp. 399–422, 2021.

[12] D. C. Nguyen, P. Cheng, M. Ding, and D. Lopez-Perez, "Enabling AI in future wireless networks: a data life cycle perspective," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 553–595, 2020.

[13] M. C. Lucic, X. Wan, H. Ghazzai, and Y. Massoud, "Leveraging intelligent transportation systems and smart vehicles using crowdsourcing: an overview," *Smart Cities*, vol. 3, no. 2, pp. 341–361, 2020.

[14] Y. Zhang, Z. Y. Dong, C. Yip, and S. Swift, "Smart campus: a user case study in Hong Kong," *IET Smart Cities*, vol. 2, no. 3, pp. 146–154, 2020.

[15] H. Mihaljevic, C. J. Larsen, S. Meier, W. Nekoto, and F. Morón Zirfas, "Privacy-centred data-driven innovation in the smart city. Exemplary use case of traffic counting," *Urban, Planning and Transport Research*, vol. 9, no. 1, pp. 426–449, 2021.

[16] B. Schneider, J. Reilly, and I. Radu, "Lowering barriers for accessing sensor data in education: lessons learned from teaching multimodal learning analytics to educators," *Journal for STEM Education Research*, vol. 3, no. 1, pp. 91–124, 2020.

[17] G. Chen, P. Wang, B. Feng, Y. Li, and D. Liu, "The framework design of smart factory in discrete manufacturing industry based on cyber-physical system," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 1, pp. 79–101, 2020.

[18] I. H. Sarker, M. M. Hoque, M. K. Uddin, and T. Alsanoosy, "Mobile data science and intelligent apps: concepts, AI-based modeling and research directions," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 285–303, 2021.

[19] C. She, R. Dong, Z. Gu et al., "Deep learning for ultra-reliable and low-latency communications in 6G networks," *IEEE Network*, vol. 34, no. 5, pp. 219–225, 2020.

[20] G. B. Rehm, S. H. Woo, X. L. Chen et al., "Leveraging IoTs and machine learning for patient diagnosis and ventilation management in the intensive care unit," *IEEE Pervasive Computing*, vol. 19, no. 3, pp. 68–78, 2020.

Hindawi

Research Article

Optimization of Product Design Mode of Art Industry Based on Multiterminal Selection Algorithm of Internet of Things

Ming Xu (ID), **Xiaoyu Ma** (ID)**, and Xuemei Liu** (ID)

School of Architecture and Art, Central South University, Changsha 410083, China

Correspondence should be addressed to Xuemei Liu; liuxm@csu.edu.cn

This article has first presented an overview and a brief design to the system, and then analyzed the system design goal and design principle. The role of the resource representation model in the Internet of Things is more important than the significance of the database schema for the web information system. In the current "vertical" application model, the device resources of the perception layer and the application entities of the business layer are in a tightly coupled relationship. In the process of system operation, it changes flexibly as the situation changes. This article will combine the characteristics of resources in the Internet of Things, from the perspectives of resources and services, and refer to the basic structure of the traditional Internet resource reference model OWL-S to appropriately expand and tailor it, and describe the relationship between Internet of Things device resources and application services. Therefore, the resource representation model oriented to the Internet of Things service in this article is proposed and simulated and verified. To a certain extent, this model can provide support for the unified description of IoT resources and services, so as to better adapt to the development needs of diversified IoT services. Based on the application and service in the Internet of Things environment restricted by time and space correlation and limited resources, the article discusses and proposes a multiterminal selection algorithm for Internet of Things business, and converts it into a mathematical combination optimization problem. Based on the ant colony optimization algorithm, a solution is given, which can realize the optimization of equipment combination scheme under resource-constrained conditions to maximize the total benefit value of the service, and can effectively improve the utilization rate of terminal resources. In the research on the analysis of the development strategy of China's industrial design industry in the era of intelligent interconnection, the analysis of the development strategy of the industrial design industry in the era of intelligent interconnection through analysis tools of industrial strategy reveals the new trend of industrial design in the era of intelligent interconnection. The analysis and formulation of industrial strategy are complicated, and it is a systematic project involving all aspects. The rational use of corresponding strategic analysis tools can further enhance the scientificity and rationality of industrial design strategy.

1. Introduction

The concept of "Internet of Things" was originally developed in the late twentieth century. The top laboratories in the field of networked RFID and emerging remote sensing technologies, the Auto-ID Laboratory of the Massachusetts Institute of Technology, automatically identify art and industrial products in the supply chain [1, 2]. With the "ITU Internet Report: Internet of Things" published by the International Telecommunication Union, the Internet of Things has been promoted to the public eye, and there has been an explosive growth in the research and application of

the Internet of Things worldwide [3]. The technical guide pointed out that the era of interconnection and interoperability of the Internet of Things is coming. Any object in real life, from toothbrushes and needles, to cars and houses, can rely on the Internet of Things technology for active interaction and data sharing [4]. With the rapid development of the Internet of Things technology, key basic technologies such as wireless electronic label technology, wireless sensor network technology, and intelligent embedded technology have matured, and the cost of related commercial equipment has fallen, and the Internet of Things has received more and more attention from more and more fields [5]. Concepts

such as smart home, smart grid, smart transportation, smart industry, and smart medical care are all concepts and technological development routes that result from combining the Internet of Things technology with this field [6].

At present, almost all IoT platform architectures are vertical application architectures; that is, they are mainly oriented to applications and closed-loop applications in the relevant fields of the industry [7]. This type of architecture has poor reusability, huge investment in one-time construction, poor data sharing and operational interaction capabilities between devices and applications, and difficulties in interconnection and intercommunication between different application systems. The main reason for this problem is the prominent problem of device heterogeneity and network heterogeneity, and the serious differentiation of IoT applications in different fields has led to the research of existing IoT applications and architectures [8]. Although with the growth of the actual application and deployment of the Internet of Things, there have been many research studies and results that combine the Internet of Things and the cloud in recent years [9, 10]. The format is uniformly regulated, and the powerful processing capabilities of cloud technologies such as cloud computing and cloud storage are used to process large data volume device data to form a backend centralized architecture system [11]. However, the research direction of the Internet of Things of this kind of system architecture is also oriented to the research of the Internet of Things technology in this field [12]. The relative equipment types are fixed and unified, and the equipment data are single. It can better deal with the problem of equipment and data unification. At the same time, this type of architecture has harsh requirements for the network and cloud servers, and a stable and reliable connection is required to obtain Internet of Things services and computing capabilities [13]. There are also certain requirements for bandwidth, and the data transmission pressure of equipment capable of carrying large amounts of data is required.

This article gives an overall overview of the Internet of Things device access system, and describes the design goals and principles of the system. From different aspects, namely, functional and nonfunctional, a detailed demand analysis of the Internet of Things device access system is carried out. In terms of the architecture design of the Internet of Things device access system, the overall design of the system is analyzed, and the functional architecture design of the system is described in accordance with the functional modules of the system. This article proposes a multiterminal selection algorithm for IoT services. Under the condition of limited resources in the Internet of Things environment, reasonably we configure the best terminal set for each user business in the business layer to improve the efficiency of the use of ubiquitous peripheral resources in the Internet of Things. A large number of users requesting services in the same time period will cause competition for equipment resources. Therefore, efficient multiterminal collaboration needs to ensure that the overall service benefit value reaches the best value, without the imbalance of user service quality. For this reason, the problem is modeled as a multiterminal selection process based on the service model and resource model. Finally, the problem is simplified into a combinatorial optimization process, and a heuristic ant colony optimization algorithm is introduced to solve the problem. Through simulation, the random allocation algorithm and the greedy algorithm are compared in terms of comprehensive benefits, utilization rate, and business success rate. Based on the diamond model to study the development of China's industrial design industry under the background of the era of intelligent interconnection, this article analyzes the production factors, demand conditions, related and supporting industries, corporate strategy, organizational structure, and factors of competition in the industry, government, and opportunities of China's industrial design industry. The industry should improve the efficiency of division and cooperation in the industry, promote competition and cooperation among enterprises, stimulate design creativity, and scientifically enhance the competitiveness of the national industrial design industry.

2. Related Work

Relevant scholars believe that the development of the Internet of Things not only requires comprehensive consideration of the security and technical issues closely related to its industry, but also needs to make a detailed future development plan for the Internet of Things industry and raise it to a strategic height [14]. The researchers pointed out that the development of core industries has become the top priority of the development of the Internet of Things industry, and it is necessary to fully support leading companies and use them to drive related industries to form a complete industrial chain [15]. Based on existing conditions, application demonstrations are mainly carried out in key areas such as strong industry driving, high relevance, and good demonstration effects, such as people's livelihood services and infrastructure. Relevant scholars believe that "the two ends are weak and the middle is strong" is the main problem in the development of China's Internet of Things [16]. The breakthrough of core technology and the realization of independent innovation are the focus of the development of the Internet of Things. Scientific and reasonable industrial layout, regional coordination and unity, avoiding duplication of construction, and focusing on regional advantages are indispensable conditions for the sound development of the Internet of Things industry.

Relevant scholars believe that it is better to regard the Internet of Things industry as an innovation system, use the method of innovation system element linkage, and build a conceptual analysis model of multi-element linkage of the Internet of Things industry innovation system from the perspective of "linkage" [17]. By analyzing the development history of the Internet of Things industry, it is concluded that the role of the government is the main innovation fulcrum of the development of the Internet of Things. Therefore, in the development of the Internet of Things, the government must create a good external environment to guide the scientific development of the Internet of Things industry. Based on the theory of value, the researchers put forward the theory of value creation of the Internet of Things and the theory of

value creation of the Internet of Things from the perspective of industrial clusters and industrial chains [18]. This theory explains the essential question of how the Internet of Things creates value as a new production tool. Relevant scholars believe that it is necessary to rely on the industrial advantages of each region to handle the relationship between traditional industries and strategic emerging industries, so as to realize the adjustment of the regional economic structure, and rely on industrial advantages to strengthen technological innovation and promote the transformation of its results [19].

Related scholars have proposed that the Internet of Things and the sensor network are actually different representations of the same thing [20]. The comparison between the networking and the sensor network and the conclusion that the two are larger and the smaller are one-sided. The researchers pointed out that the ubiquitous integrated sensors, the subtle nodes of the communication unit, and the data processing unit form a wireless network in a self-organizing manner [21]. The items are given "wisdom" by these subtle details, and they interact and influence each other according to the instructions received, so as to achieve the integration of human economic society and the physical world. Based on the above viewpoints, it can be known that the Internet of Things is defined mainly based on two technologies in the research of Chinese scholars, namely, radio frequency identification technology and sensor network technology.

Related scholars' interpretation of the concept of the Internet of Things includes the above two theories [22]. He pointed out that the Internet of Things is an intelligent sensor that places radio frequency identifiers in objects that need to be monitored, and then recognizes GPS and other technologies, and communicates with the Internet through interfaces. The definition of the concept of the Internet of Things by domestic and foreign scholars will vary according to the focus, that is, the core technology involved [23]. But no matter from which point of view it is defined, the core connotation is basically the same; that is, the Internet of Things is a combination of the Internet and information sensing equipment. Through the construction of an intelligent network connected to things, communication and information exchange in an agreed manner are finally realized.

Relevant scholars have summarized the application fields of the Internet of Things technology, including industrial automation, smart agriculture, pollution prevention, public management, medical and health, communication logistics, smart home, and other major fields [24]. The changes brought about by the Internet of Things in the four fields of agriculture are described in detail. Relevant scholars looked forward to the application of the Internet of Things technology in agriculture and pointed out that the application of the Internet of Things technology to agricultural production in China, a large agricultural country, is a good opportunity to solve the "three rural" problems. Relevant scholars believe that the application of the Internet of Things has involved food traceability, industrial monitoring, intelligent transportation, environmental protection, environmental

monitoring, government work, public safety, enemy investigation and intelligence collection, smart fire protection, safe home furnishing, and water system monitoring [25]. It can be seen that most domestic scholars are studying the application of the Internet of Things in a wider range of industries, but they have not made further detailed and indepth studies on the specific application of the Internet of Things in these industries. The common characteristics of the application have not been clearly mentioned.

3. IoT Device Access System Design

3.1. System Requirements Analysis. The smart building monitoring system collects information such as temperature, air conditioning location, and carbon dioxide concentration through the underlying sensor equipment to achieve comfortable management of the working environment and improve work efficiency. In IoT applications, there are a large number of underlying sensor devices, and the transmission protocols are different and diverse. The IoT device access system provides a system for the access management of underlying sensor devices and provides the service of encrypting data before transmission.

The topic proposed in this article is based on the IoT smart IoT sensor equipment project. In this project, the IoT equipment is connected to the system and is responsible for the management of the sensor equipment at the lower level perception layer, and provides services such as data transmission to the upper layer. Figure 1 shows the architecture of the smart IoT sensor device. The underlying sensor equipment is composed of traditional IoT sensor equipment plus NB modules. After the IoT equipment is connected to the system, it will be managed by the system, and the collected information usage and other data will be safe and complete. It is transmitted to the upper application server, and the data are processed and then visualized into the intelligent information management system.

When uploading data to the system, the system will select the corresponding protocol according to the transmission protocol used and then parse the original protocol data to obtain intuitive environment perception data, and re-encapsulate the data into a unified data format for transmission to the system. In the process of transmission, in order to ensure the security of data transmission, the data are encrypted and then transmitted, and the upper management system can obtain the data completely.

3.2. Demand Analysis of IoT Device Access System. The underlying sensor device network layer has the characteristics of multiple types of devices, limited computing and storage capabilities of the device, and so on. Therefore, for the communication protocol, the corresponding protocol should be used according to the characteristics of the underlying sensor device layer to ensure the accuracy and fluency of data transmission. The Internet of Things device access system has functional modules such as user management, application product management, device management, protocol management, data analysis and
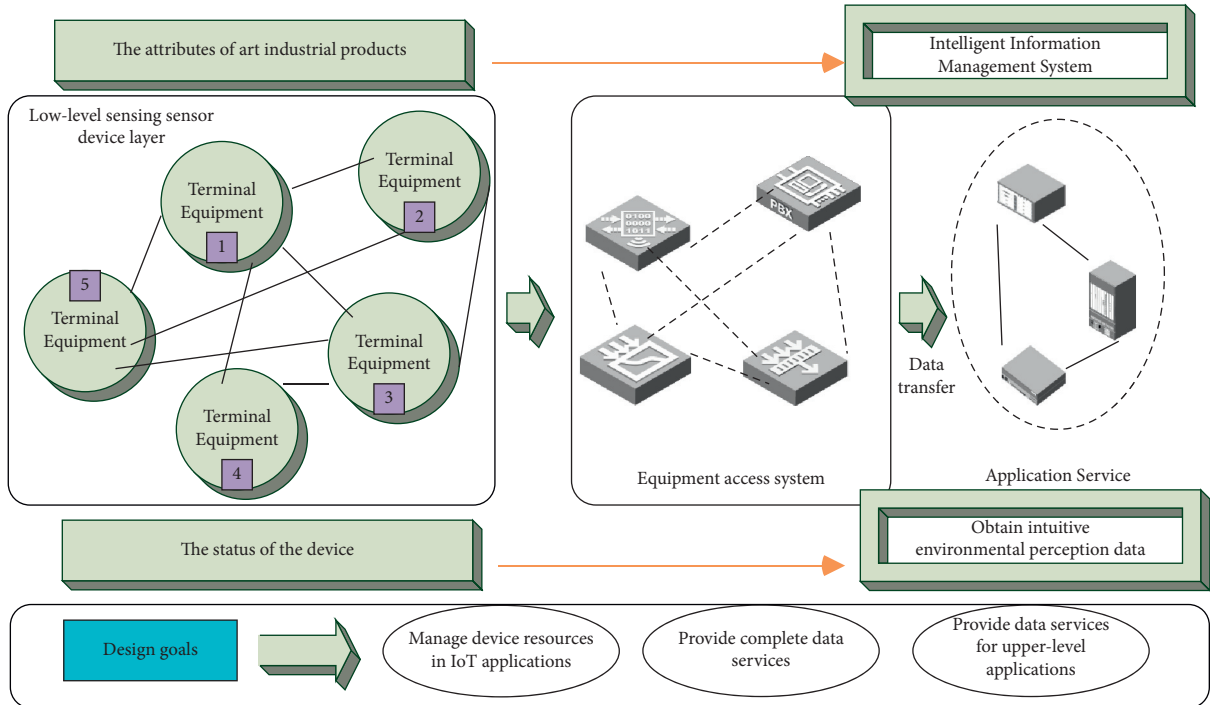
Figure 1: Smart IoT sensor device architecture.

encapsulation, and data encryption transmission. The specific requirements are analyzed as follows:

(1) User information management is mainly responsible for managing the basic information stored by users when they use IoT devices to access the system. It has the following basic functions: user registration and login—this function is responsible for entering the system; and password modification and retrieval—this function is responsible for the management of user passwords. In addition, the management of user information also contributes to the management of modules such as application product information and equipment.

(2) The basic information of the Internet of Things device access system management application products mainly includes the creation of application products, the modification of application product information, the deletion of application products, and the query of application product information. Moreover, the management of application product information also contributes to the management of equipment modules.

(3) In IoT applications, different application requirements lead to different sensor devices. In the IoT device access system, devices are managed according to their application information. In an IoT application, there will be multiple sensor devices serving it. Multiple sensor devices can be saved under the same application product information. The equipment management module is based on the management of application product information and is responsible for functions such as adding and modifying equipment attribute

information, deleting equipment, and querying equipment information.

(4) Different sensor equipment manufacturers will lead to different communication protocols supported by the equipment. The Internet of Things equipment access system supports mainstream communication protocols, such as HTTP protocol, MQTT protocol, and CoAP protocol, so as to support multiple sensor equipment.

(5) After the IoT device access system collects the original data from the sensor device network of the underlying perception layer, it will parse the original data according to the data transmission protocol supported by the sensor device and then encapsulate it into a unified, standardized, and intuitive data. When the data are transmitted, it is encrypted and then uploaded to ensure the accuracy and integrity of the data.

(6) The Internet of Things device access system stores the user's information, the information of the Internet of Things application products and the information of the sensor equipment in the database, so that the user can modify, query, and delete these data information.

3.3. Design of the Internet of Things Device Access System Architecture. The Internet of Things device access system is an important part of the entire Internet of Things applications. It solves some of the problems that will be involved in the development of Internet of Things applications, such as the access management of multiple different resource devices, and the analysis and packaging of raw data. Figure 2
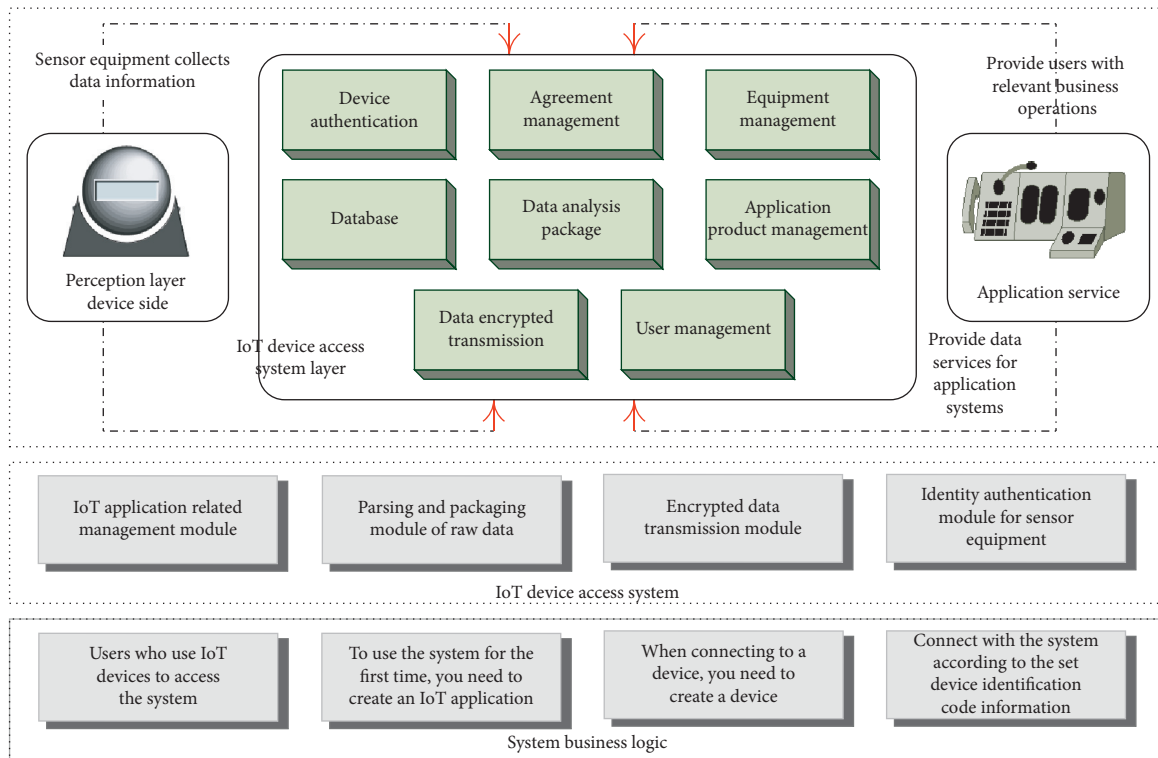
FIGURE 2: Functional architecture diagram of IoT device access system.

shows the overall architecture of the Internet of Things device access system.

In this architecture, sensor devices are responsible for collecting data information. In the application scenario of this topic, the data information related to the IoT sensor equipment are collected, which is the most basic and core part of the IoT application implementation process. Wireless communication technology is mainly used to realize the collection of related information and other operations. The IoT device access system layer is mainly responsible for providing data for the upper layer, combining with database technology to manage the sensor devices in the underlying perception layer, and adapting appropriate IoT application protocols to parse the collected raw data. The application server is mainly responsible for providing data services for the upper-layer application system, and parses and processes the received lower-layer data for use by the upper-layer business system.

4. Multiterminal Selection Algorithm of Artistic Industrial Products for IoT Sensor Devices

4.1. Problem Description. The execution of ubiquitous services of the Internet of Things often requires the collaborative combination of multiple terminals on the perception layer to achieve. Because of the differences and differences in the configuration of application scenarios, resources, etc., compared with traditional Web services, IoT services have very different characteristics, such as heterogeneity and the execution of services are restricted by device resource conditions. It is precisely because of these different characteristics that the collaborative combination of multiple terminals in the ubiquitous environment of the Internet of Things faces the following challenges:

(1) Since devices and network links are shared in a ubiquitous environment, it is a more complicated problem to configure corresponding device sets and link sets for multiple ubiquitous services in a bandwidth-constrained network reasonably and efficiently.

(2) Due to the influence of mobility and other aspects, the state of terminal equipment in the ubiquitous environment of the Internet of Things is usually unstable, such as processing delay, equipment reliability, and other availability of equipment. Therefore, periodic dynamic adjustments are required to meet user requirements for service quality.

Due to the limited network resources and terminal capabilities, a large number of users requesting services at the same time period will lead to equipment resource competition. Therefore, efficient multiterminal collaboration needs to ensure that the overall service benefit value is optimal, and there should be no imbalance in user service quality. Multiterminal selection should focus on solving the problem of how to maximize the sum of the benefits of business-layer service requests under the condition of limited underlying perception resources and network resources, so as to improve the use efficiency of terminal equipment and network links.

Most of the mapping models between the business layer and the resource layer in the multiterminal collaborative

aggregation mechanism in the ubiquitous environment build constraint functions from the perspective of resource constraints. There are relatively few studies on the abstract modeling work of IoT business layer and resource layer.

To this end, this article will first abstract and model the IoT business layer (service model) (SM) and the perception resource layer (resource model) (RM), and establish a multiterminal selection management layer between the network layer and the business request layer (resources aggregation management layer) (RAML); the description information of the underlying resources will be automatically mapped to the terminal aggregation management layer. This layer will realize the combination and allocation of resources based on the multiterminal selection algorithm ACO-MTS for ubiquitous services of the Internet of Things proposed in this article to ensure the maximum and smoothness of comprehensive benefits. In order to adapt to the heterogeneity and instability of terminals in the ubiquitous environment of the Internet of Things, a comprehensive benefit function based on the maximum business success rate and benefit is designed. Finally, through experimental simulation and random selection (RS) algorithm, the comprehensive benefits, the total consumption of resources, and the effective utilization of resources are compared to show the effectiveness of the mechanism.

4.2. Mathematical Model of Multiterminal Selection Problem.

The resource layer is a comprehensive modeling of the network layer and the ubiquitous perception layer in the overall architecture of the Internet of Things. The models are all represented in the form of logic diagrams, which are, respectively, represented as GSM and GRM, and the problem to be studied for multiterminal selection for IoT services is how to map GSM to GRM to select the best terminal set to provide services.

TAS means that for each ubiquitous service i, there are multiple optional terminal collaboration sets corresponding to it, denoted as $C_{i,s}$, where s represents the s-th terminal collaboration set corresponding to the ubiquitous service Pi. At the same time, the TAS corresponding to each ubiquitous service must meet the corresponding resource capacity constraints, mainly including terminal capacity and link capacity limitations, where w represents the type of terminal k's capacity, and e represents which one is occupied link:

$$\prod_i T_{kw} > t_{i,kw} \quad w \longrightarrow (0, W),$$
$$\prod_i L_e > l_{i,e} \quad e \longrightarrow (0, E). \tag{1}$$

Combining the above concepts of business model, resource model, and terminal collaborative aggregation model, it can be seen that the key to the multiterminal selection problem for IoT services lies in the SM to RM mapping algorithm in the resource aggregation management layer (RAML). RAML is responsible for multiple universal applications at the application layer.

4.3. Selection Algorithm Analysis.

In order to further abstract the mathematical model of the mapping from SM to RM, the benefit function is mainly measured by business quality here. Assuming that each ubiquitous business i corresponds to a total of Si TAS, in order to avoid the supervisor dependence of the benefit evaluation function, the TOPSIS criterion of the multi-attribute decision-making scheme is also introduced to calculate the benefit index of each TAS.

After determining the evaluation criteria of the benefit function, the Multi-Terminal Selection Problem (MTSP) can be expressed as

$$Max \quad \prod_{i=0}^{l-1} \prod_{s=0}^{S_i-1} U(C_{is}, i) = \prod_{i=0}^{l-1} \prod_{s=0}^{S_i-1} (x_{is}\varphi_{is}),$$

$$s.t. \begin{cases} \prod_{i=0}^{I-1} \prod_{s=0}^{S_i-1} \prod_{e=0}^{E-1} z_{i,se} l_{i,e} > L_j, \qquad j \longrightarrow (-1, J), \\ \prod_{i=0}^{I-1} \prod_{s=0}^{S_i-1} \prod_{v=0}^{V_i-1} y_{i,svk} t_{i,vw} > T_{kw}, \quad w \longrightarrow (0, W) \cup k \longrightarrow (-1, K). \end{cases} \tag{2}$$

From the analysis of the above formula, it can be seen that the MTSP problem is an integer linear programming problem with multidimensional constraints. Due to the difficulty and complexity of solving this type of combinatorial optimization problem, in order to simplify the solution process, it is necessary to reduce the dimensions of the multidimensional constraints in the formula, that is, to convert multiple limited constraints into unique constraints. Therefore, the MTSP problem can be transformed into a Multi-dimension Multi-choice Knapsack Problem (MMKP) through dimensionality reduction processing; that is, the two constraints on terminal capacity and link capacity in MTSP are combined into one-dimensional constraints. The conversion formula is

$$r_{i,sk} = \begin{cases} \prod_{e=0}^{E-1} (k' z_{se} l_{i,e}), \qquad k' \in [KW - 1, KW + J], \\ \prod_{v=0}^{V_i} (k' t_{i,v} y_{i,sv}/w), \quad k' \in [-1, KW - 1], \end{cases} \tag{3}$$

$$R_{k'} = \begin{cases} L_{KW+k'}, \quad k' \in [KW - 1, KW + J], \\ R_{k'/W}, \quad k' \in [-1, KW - 1]. \end{cases}$$

Thus, the MTSP problem can be transformed into a one-dimensional constrained MMKP mathematical model:

$$Max \quad \prod_{i=0}^{I-1} \prod_{s=0}^{S_i-1} U(C_{is}, i) = \prod_{i=0}^{I-1} \prod_{s=0}^{S_i-1} (x_{is} \bullet \varphi_{is}),$$

$$s.t. \begin{cases} \prod_{s=0}^{S_i-1} x_{is} = 1, \qquad i \longrightarrow [1, KW], \\ \prod_{i=0}^{I-1} \prod_{s=0}^{S_i-1} x_{is} r_{vk'} < R_{k'}, \quad k' \longrightarrow [-1, KW + J]. \end{cases} \tag{4}$$

MMKP is an NP-hard problem with a broad engineering background. Many practical application problems can be described as MMKP models, such as inventory compression problems and distributed computing system processor allocation strategy problems. Here, the ant colony optimization (ACO) algorithm is introduced to solve the MMKP problem, and the ant colony algorithm model is appropriately modified according to the characteristics of the multiterminal selection problem model to form the ACO-MTS algorithm to adapt it to the MMKP problem model.

When the ACO algorithm is applied in different scenarios, the pheromone update model in the multiterminal selection algorithm is often different. The pheromone update process of each TAS of the ACO-MTS algorithm is shown in the following formula:

$$\tau_s(t-n) = \prod_{k=0}^{K-1} \tau_{k,s} + (1+p)\tau_s(t-2n), \quad (5)$$

$$s.t. \quad \tau_{k,s} = Q\varphi_{is}/C^k,$$

where Q is a constant and can be set to 1, $\triangle\tau$ is the pheromone increment produced by the k-th ant passing through the s-th TAS terminal set, p represents the volatilization degree of the pheromone, and C^k represents the total benefit value of all TAS selected by the k-th ant. It can be seen from the above formula that the higher the benefit value of TAS, the larger the corresponding pheromone increment. At the next iteration, all ants will calculate the probability of each TAS being selected based on the latest pheromone vector. The calculation method is

$$p_s(k) = \begin{cases} \tau_{a,s}(t)\eta_{\beta,s}(-t), & s \subset tabu(k), \\ -1, & s \not\subset tabu(k). \end{cases} \quad (6)$$

In the dynamic ubiquitous peripheral environment of the Internet of Things, due to the limited resource capabilities of devices such as terminals and links in the network, it is often impossible to undertake all the requested ubiquitous services at the same time. However, there are many types of services in the Internet of Things, with outstanding heterogeneity characteristics, and different services have different requirements for indicators such as delay, bit error rate, and packet loss rate.

5. Analysis of the Development of Industry Models

5.1. The Impact of the Era of Intelligent Interconnection on the Development of Art and Industrial Products. Products will no longer be a combination of engineering structure and CMF, and more and more products are evolving into smart terminals. When we design a new product, it is far from enough to only consider the ergonomics and practical functions of the past. What we are facing is the features and capabilities of computing power, smart application, information interface, interconnection, and cross-border integration. Moreover, they often have a powerful data backend, and continue to provide users with updated and iterative service content. On the other hand, consumers have been integrated with

intelligence and services. Users put forward beautiful experience and simple and easy-to-use expectations for complex smart products, and it is an important task of design to transform complex scientific logic into a beautiful experience. Figure 3 shows the comparison chart of total business benefits.

Intelligent design has been widely and deeply applied in the field of visual communication. Giants such as Alibaba, Jingdong, Google, and Amazon have developed their own computing designs. Among them, the intelligent design robots of Alibaba and JD.com have made great progress in intelligent analysis of user data, automatic generation of pictures, and intelligent push based on user big data. In the field of product design, the integration of user big data and design modules, technical parameters, collaborative platforms, etc., intelligent design is still in its infancy. The integration of IoT and industrial design forms an innovative model of intelligent design.

Intelligent manufacturing puts forward many new requirements for industrial design. Industrial design is changing from the optimization of a single product to the full life cycle of design, production, management, and service. Industrial design is closely related to intelligent manufacturing.

5.2. The Important Role of the Transformation and Upgrading of the Industrial Design Industry in the "Dual Cycle" Strategy. The economy and society are in a dynamic and circular process, and the various elements are closely connected. Only when the elements are communicated smoothly can the economy develop healthily. In view of the current stage of economic development and the internal and external environment of economic development, the Party Central Committee proposed a development situation where domestic economic cycles are the mainstay, and domestic and foreign economic cycles are supplemented by mutual promotion. Therefore, the formulation of industrial design industry development strategies should firmly grasp the "double cycle." Under the long-term downturn of the world economy, grasp the main characteristics of the current turbulent external environment, rising protectionism, and shrinking international markets. Speeding up the "dual cycle" strategy is an important measure for China to stabilize the domestic economic environment and seek development opportunities. The "inner loop" shows China's market size and absolute advantages. Second, the integrity of China's industrial system is not available in other countries. Therefore, expanding domestic demand, increasing domestic market demand, and focusing development on the domestic market can be better. Through the "dual cycle" strategy, China's relative advantages in cooperation and competition with other countries will be enhanced, and domestic economic development will be promoted. Figure 4 shows the comparison of the role of the transformation and upgrading of the industrial design industry in the "dual cycle" strategy.

From the perspective of international circulation, excellent industrial design can enhance the international
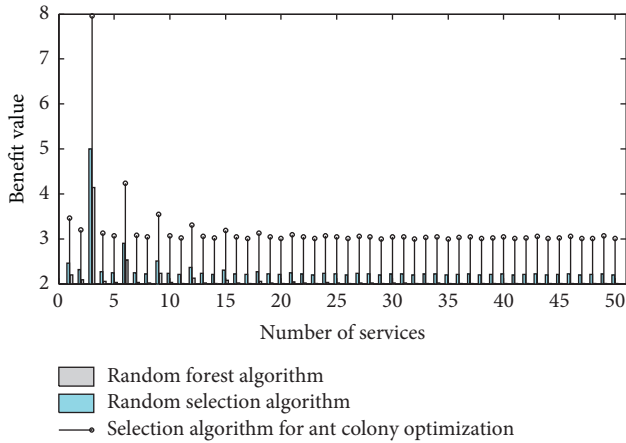
FIGURE 3: Comparison of total business benefits.
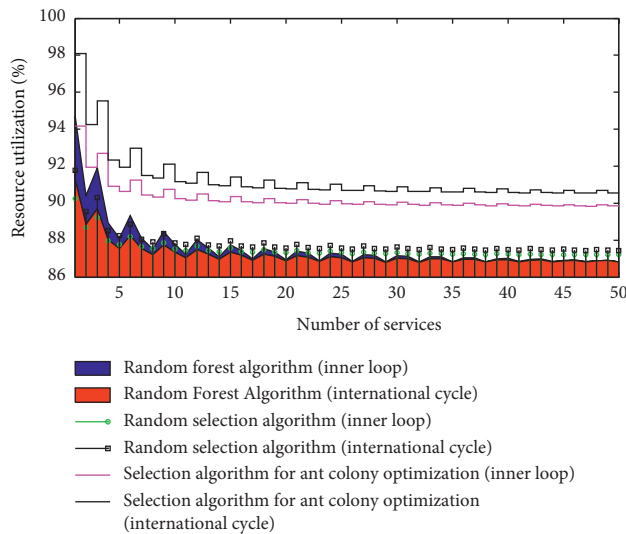
FIGURE 4: Comparison of the role of the transformation and upgrading of the industrial design industry in the "dual cycle" strategy.

competitiveness of domestic independent products, seize a place in the global market, help to promote the development of national innovation level, and enhance national competitiveness. The scale advantage of the industrial design industry has been formed, but there is still huge room for improvement in the quality advantage. Therefore, in order to further enhance China's position in the international industrial chain and ensure the effective circulation of foreign economies in the "postepidemic" era, developing the industrial design industry is an important focus for the development of China's industrial industry.

5.3. The Important Role of the Transformation and Upgrading of the Industrial Design Industry in the Supply-Side and Demand-Side Upgrading.
From the supply side, in order to adapt to the changes in people's consumption concepts, it is necessary to enhance the status of industrial design in

China's economic production. Industrial design can endow products with new added value in consumption upgrades to meet growing consumer demand. Improving the level of industrial design can promote the innovation level of the manufacturing industry. Only continuous innovation can guarantee the competitive advantage of Chinese manufacturing and increase the added value of products. It is not only necessary to increase design investment, but also to focus on the cultivation of design innovation talents, promote design investment, and protect design patents. Figure 5 shows the comparison between the transformation and upgrading of the industrial design industry in the supply-side and demand-side upgrading.

From the demand side, as people's living standards continue to improve today, the transformation and upgrading of the design industry is a huge driving force for the release of effective demand and can promote demand-side changes. With the increasing material and spiritual needs of people, the reform of consumption structure, people are pursuing a high-quality life, so that the design industry has a huge room for development. In order to meet the huge consumer market demand and provide high-quality products and services, continuous innovation is required. While meeting the needs of consumers, it has also stimulated the optimization and upgrading of the design industry, making capital continue to flow to the design innovation industry. The design innovation industry radiates to other industries to increase the added value of other industries and promote the overall innovation and upgrading of China's industrial system. The design industry not only meets consumers' growing material needs, but also meets their spiritual needs. The design industry provides consumers with differentiated and personalized services, while giving consumers a sense of identity, which is a diversification of society and the importance of social relations.

5.4. Analysis Based on the Diamond Model.
Industrial design enterprises are divided into enterprises that provide industrial design services to enterprises that provide industrial design services to industries based on production targets. The former is to directly create economic benefits for enterprises through resource integration, sharing platforms, and technical services, such as professional product design and interaction design enterprises; the latter is to provide design services for enterprises, industries, or industries at the same time, as shown in Table 1.

In the era of intelligent interconnection, China's industrial relations and economic structure have undergone drastic changes. Under the new demand and new background of domestic manufacturing development, China's industrial design industry is also facing new plans for industrial structure adjustment and industrial development. In the era of smart interconnection, industrial design will be injected with digital technologies such as smart technology and big data. Industrial design includes the characteristics of innovation and creativity-intensive,
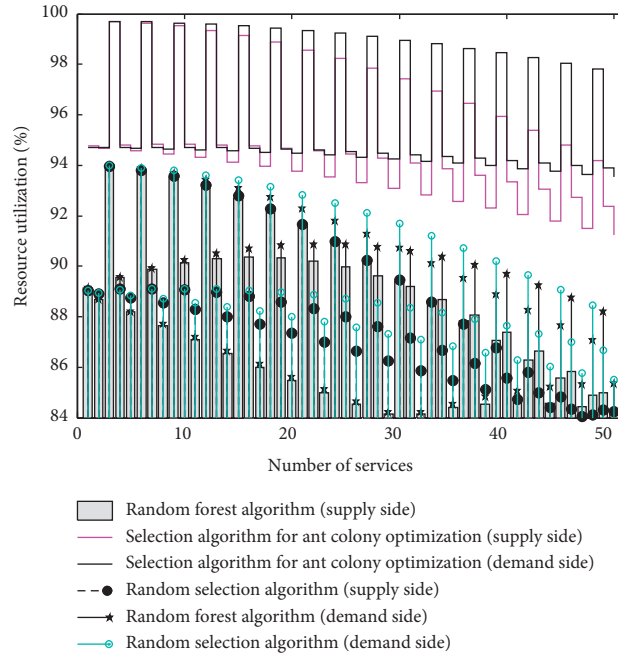
Figure 5: Comparison of the role of industrial design industry transformation and upgrading in supply-side and demand-side upgrading.

Table 1: Classification of industrial design companies.

Classification based on production objects	Classification based on the main body of production	According to the production content as the standard classification
Companies that provide industrial design services for companies	Industrial design service company	An enterprise centered on physical product design
Companies that provide industrial design services for industries and industries	Industrial design collaboration enterprise	An enterprise focusing on information product design
Art industrial product industry model	Industrial design application enterprise	An enterprise centered on product service system design

knowledge-intensive, and other characteristics. The development of industrial design has a powerful role in promoting innovation in the manufacturing industry and the entire country. Based on the diamond model to analyze China's industrial design industry in the era of intelligent interconnection, it can fully position and analyze the development of the industry, make a long-term plan for the healthy development of China's industrial design industry, and highlight the competitive advantage of China's industrial design industry in the world. The factor analysis of China's industrial design industry based on the diamond model is shown in Figure 6.

Internet technologies such as the Internet of Things and big data are widely used, making people's lives more convenient, improving people's work efficiency, and changing people's lifestyles. Under the background of this era, the application of industrial design is particularly important. Through industrial design, high-end technology can become more grounded. At the same time, intelligent technology also promotes the development of industrial design. The intervention of the Internet of Things has changed the way of working in traditional industrial design. Therefore, the industrial design industry should make

corresponding adjustments and upgrades to meet the new demands of consumers in the era of intelligent interconnection. The demand condition analysis of the industrial design industry in the era of intelligent interconnection is shown in Figure 7.

With the development of social economy, the industrial design industry chain has been gradually improved, and related supporting industries and policy planning have also been further improved. Industrial design has gradually formed a good industrial environment. At present, an industrial design industry composed of upstream industries, industrial design industries, and downstream industries has been formed. Related and supporting industries are shown in Figure 8.

Industrial design consulting companies began to transform to "design house" and other models. This model first appeared in the mobile phone industry. They cooperated with software and hardware suppliers and did not require mass production, and paid more attention to development efficiency. They can develop their own conceptual mobile phones, carry out prototype tests, and modify plans according to customer needs to produce small batches in order to cope with the current market situation with urgent
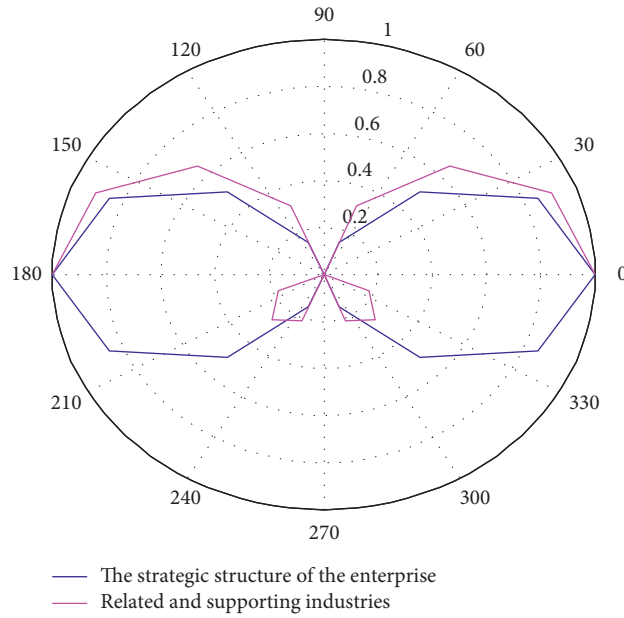
Figure 6: Analysis of China's industrial design industry factors based on the diamond model.
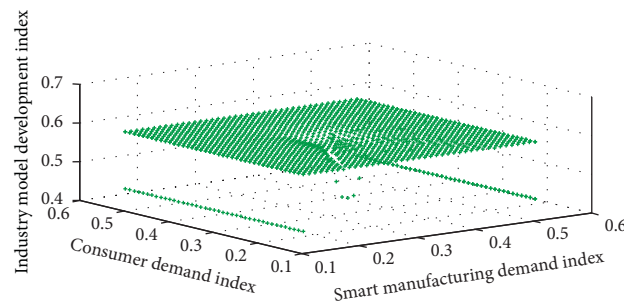
Figure 7: Analysis of the demand conditions of the industrial design industry in the era of intelligent interconnection.
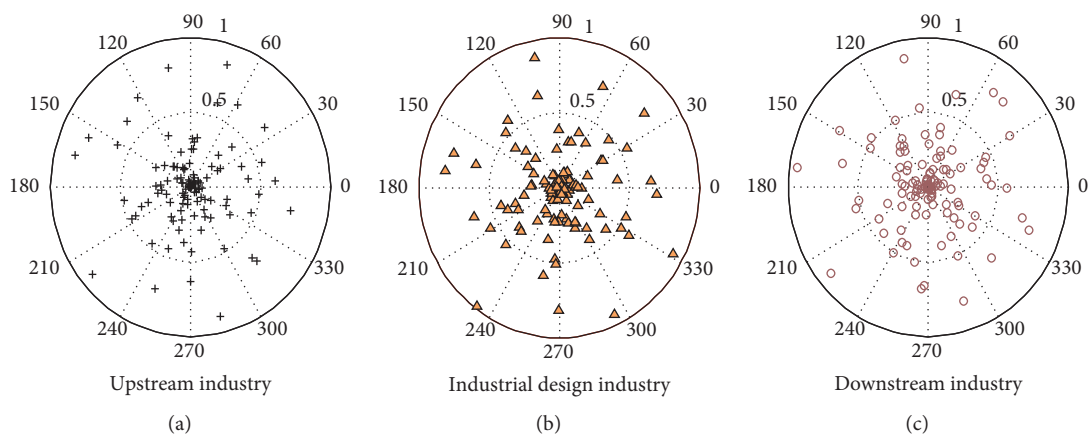
Figure 8: Analysis of related supporting industries in the industrial design industry. (a) Upstream industry. (b) Industrial design industry. (c) Downstream industry.

needs at the time and the shortcomings of domestic enterprises with low R&D capabilities. Industrial design has developed rapidly, and enterprise-level industrial design centers have gradually been established. In addition, these industrial design centers will also provide related design services externally (within the industry).

6. Conclusion

In this paper, the requirements of IoTs device access system are analyzed in detail from functional and nonfunctional perspectives, and the software system is optimized. Some principles that need to be followed in the design have been designed for the system. Based on the dimensionality reduction processing and ant colony optimization algorithm, a multiterminal selection algorithm ACO-MTS for Internet of Things business is proposed, and it is compared with the random allocation algorithm (RS) and greedy in terms of total business benefits, resource utilization, and business success rate. Compared with GA, the algorithm is improved. However, because the simulation experiment was carried out under several premises and assumptions, it failed to simulate the multiservice batch processing scenario, which still has a gap with the actual IoT multiservice request environment; at the same time, the number of evaluation indicators of the benefit function is relatively less. The next step will be to strengthen simulation experiments for multiservice batch processing scenarios and at the same time design more complex benefit functions to lay the foundation for the optimal allocation of resources in the ubiquitous network peripheral environment of the Internet of Things. The study found that the attribute of industrial design purely as a productive service industry has changed. A large number of designers and design companies are transforming from traditional design services to the design industry under the new situation and new opportunities analyzed and researched in the previous article, that is, from selling ideas at a low price. And the solution to the manufacturer is transformed into the integration of industry chain resources through the information network, creating a better and high-quality life for users, and realizing their own value at the same time.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgments

References

[1] B. Maag, Z. Zhou, and L. Thiele, "A survey on sensor calibration in air pollution monitoring deployments," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4857–4870, 2018.

[2] H. Habibzadeh, Z. Qin, T. Soyata, and B. Kantarci, "Large-scale distributed dedicated- and non-dedicated smart city sensing systems," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7649–7658, Dec. 2017.

[3] K. Kim, J. Lee, and J. Choi, "Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system," *IEEE Communications Letters*, vol. 22, no. 4, pp. 828–831, Apr. 2018.

[4] K. W. Choi, P. A. Rosyady, L. Ginting, A. A. Aziz, D. Setiawan, and D. I. Kim, "Theory and experiment for wireless-powered sensor networks: how to keep sensors alive," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 430–444, Jan. 2018.

[5] A. W. Al-Dabbagh and T. Chen, "Sounding off on industrial alarm systems," *IEEE Potentials*, vol. 37, no. 2, pp. 24–28, Mar. 2018.

[6] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, "Next-generation big data analytics: State of the art, challenges, and future research topics," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1891–1899, Aug. 2017.

[7] Y. Yang, Z. Zheng, K. Bian, L. Song, and Z. Han, "Real-time profiling of fine-grained air quality index distribution using UAV sensing," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 186–198, Feb. 2018.

[8] J.-H. Suh, J.-H. Park, S.-J. Kweon, and H.-J. Yoo, "Multiparameter sensor interface circuit with integrative baseline/offset compensation by switched-capacitor level shifting/balancing," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 3, pp. 316–320, Mar. 2018.

[9] M. Taki, S. Abdanan Mehdizadeh, A. Rohani, M. Rahnama, and M. Rahmati-Joneidabad, "Applied machine learning in greenhouse simulation; new application and analysis," *Information Processing in Agriculture*, vol. 5, no. 2, pp. 253–268, 2018.

[10] I. Qasim, M. W. Anwar, F. Azam, H. Tufail, W. H. Butt, and M. N. Zafar, "A model-driven mobile HMI framework (MMHF) for industrial control systems," *IEEE Access*, vol. 8, pp. 10827–10846, 2020.

[11] P. Basanta-Val, "An efficient industrial big-data engine," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1361–1369, Apr. 2018.

[12] M. Chen, J. Yang, L. Hu, M. S. Hossain, and G. Muhammad, "Urban healthcare big data system based on Crowdsourced and cloud-based air quality indicators," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 14–20, Nov. 2018.

[13] S. Choi, D. J. Kim, Y. Y. Choi et al., "A multisensor mobile interface for industrial environment and healthcare monitoring," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2344–2352, Mar. 2017.

[14] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

[15] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4798–4810, Sep. 2014.

[16] M. W. Anwar, M. Rashid, F. Azam, A. Naeem, M. Kashif, and W. H. Butt, "A unified model-based framework for the simplified execution of static and dynamic assertion-based verification," *IEEE Access*, vol. 8, pp. 104407–104431, 2020.

[17] J. Zhu, Z. Ge, and Z. Song, "Distributed parallel PCA for modeling and monitoring of large-scale plant-wide processes with big data," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1877–1885, Aug. 2017.

[18] G. Mois, S. Folea, and T. Sanislav, "Analysis of three IoT-based wireless sensors for environmental monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 2056–2064, Aug. 2017.

[19] L. Li, K. Ota, and M. Dong, "When weather matters: IoT-based electrical load forecasting for smart grid," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 46–51, Oct. 2017.

[20] W. Zou, F. Yao, B. Zhang, C. He, and Z. Guan, "Verification and predicting temperature and humidity in a solar greenhouse based on convex bidirectional extreme learning machine algorithm," *Neurocomputing*, vol. 249, pp. 72–85, Aug. 2017.

[21] M. Bockelmann, P. Nickel, and F. Nachreiner, "Development of an online checklist for the assessment of alarm systems and alarm management in process control," *Lecture Notes in Computer Science*, vol. 10294, pp. 325–332, 2017.

[22] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.

[23] S. C. Folea and G. D. Mois, "Lessons learned from the development of wireless environmental sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3470–3480, Jun. 2020.

[24] Z. Y. Wu and A. Rahman, "Optimized deep learning framework for water distribution data-driven modeling," *Procedia Engineering*, vol. 186, pp. 261–268, Jan. 2017.

[25] M. Bockelmann, P. Nickel, and F. Nachreiner, "The design of alarm systems and alarm management - an empirical investigation from an ergonomic perspective," *Advances in Intelligent Systems and Computing*, vol. 604, pp. 507–515, 2017.

Research Article

The Application of Automatic Identification System Information and PSO-LSTM Neural Network in CRI Prediction

Wei Zhou [1,2] **Yun Li,**[1,2] **Yingjie Xiao,**[1,2] **and Jian Zheng**[3]

[1]*Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China*
[2]*Shanghai High-level Local University Innovation Team (Maritime Safety & Technical Support), Shanghai 201306, China*
[3]*College of Transport and Communications, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Wei Zhou; weizhou@shmtu.edu.cn

Considering that collision accidents happen sometimes, it is necessary to predict the collision risk to ensure navigation safety. With the information construction in maritime and the popularity of automatic identification system application, it is more convenient to obtain ship navigation dynamics. How to obtain ship encounter dynamic parameters through automatic identification system information, assess ship collision risk, find out dangerous target ships, and give early warning and guarantee for ship navigation safety, is a problem that scholars have been studying. As an index to measure the degree of ship collision risk, CRI, namely, collision risk index, is usually obtained by calculating ship encounter parameters and comprehensive analysis. There are many factors that affect CRI, and the values of many parameters depend on expert judgment. The corresponding CRI has nonlinear and complex characteristics, which is highly correlated with the time sequence. In order to enhance the prediction accuracy and efficiency, PSO-LSTM neural network is applied in the paper to predict CRI. Experiments show that PSO-LSTM neural network can effectively predict collision risk and provide a reference for navigation safety.

1. Introduction

Ship collision accidents are the main type of marine accident, and are more likely to occur in port areas with dense ship traffic. It is usually accompanied by casualties, property losses and marine environmental pollution, which often attracts extensive attention. For example, according to Reference [1], at approximately pm 7:50, Jan.6th, 2018, the Panamanian tanker "Sanchi" collided with the bulk carrier "CF CRYSTAL," resulting in ship sinking, casualties and oil pollution. Due to the serious consequences, people pay special attention to ship collision, which is assessed, forecast and forewarning, so as to avoid the occurrence of collision accidents and reduce the related losses. Some scholars have proposed the ship collision risk index (CRI), which is widely used to quantify the degree of collision risk [2]. Generally, parameters such as DCPA and TCPA are used to quantify ship collision risk [3–5]. Then, the ship domain, which is a safe area around the own ship, is applied to measure the

collision risk [6–9]. In Reference [10], once the target ship enters the own ship's domain, there may be collision risk. With the popularity of AIS data and the convenience of data acquisition, more and more new methods have been used to evaluate ship collision risk. In References [11, 12], AIS data are used to measure the risk level of ship collision in the waters. Reference [13] uses evidence theory to evaluate ship collision analysis. According to Reference [14], the artificial neural network method is used to quantify ship collision analysis. Fuzzy inference method is used to calculate ship collision risk by Reference [15]. In Reference [16], a quantitative assessment algorithm based on SVM technology can obtain collision risk according to ship sailing status. From Reference [17], images are constructed by AIS data, and then a method based on CNN and image recognition is used to classify ship collision risks. In short, with the trend of intelligent navigation, more and more intelligent technologies have begun to be applied in the field of marine transportation to serve and ensure the safety of ship

navigation. Considering that the ship driver not only needs to know the current collision risk of the ship, but also wants to predict the possible risk, it is necessary to predict the collision risk. Taking an example as follows: reference [18] deduces an ordered probabilistic regression model to study perceived collision risk. In Reference [19], the domain method is adopted to carry out dynamic risk warning for ship encounter process.

LSTM neural network, as an excellent deep learning model, has been widely used. An automatic collision avoidance algorithm based on LSTM for ship continuous motion space is proposed in reference [20], and the experiment shows that it is effective. PSO can optimize hyperparameters of LSTM. Therefore, PSO-LSTM neural network can be used for more accurate prediction. For instance, according to reference [21], ship traffic flow prediction is currently faced with problems of high randomness, many influencing factors and low accuracy, so PSO-LSTM prediction model is proposed to improve the prediction accuracy. In reference [22], due to the high randomness and complexity of ships sailing at sea, it is difficult to make accurate prediction. PSO-LSTM is applied in forecast ship's comprehensive posture. Considering that the measurement of collision risk depends on expert judgment, the ship collision risk index has the characteristics of nonlinear, random, empirical, fuzzy and complex, etc. Meanwhile, as ship encounter process parameters are a time sequence with correlation, this paper proposes using PSO-LSTM model for prediction analysis.

The main contribution of this paper is as follows: PSO-LSTM model is applied to predict collision risk. The feasibility of the prediction method is analyzed through the design of single-ship and multi-ship encounter situation experiments. The change characteristics of CRI during ship encounter are studied and visualized. AIS data is collected by ship maneuverability simulator test to ensure the correctness, integrity and synchronization of data, so as to ensure the accuracy of CRI prediction data sources. By comparing with pilot's experience in BRM practical operation test, the conformity of CRI and its prediction with the actual situation is verified.

The rest of this paper is organized as follows: in Section 2, the quantitative methods of collision risk are discussed, including ship collision parameters and fuzzy membership quantification methods. In Section 3, the basic principles of PSO algorithm and LSTM model are introduced, and the ship collision risk prediction model using PSO-LSTM is established. Simulation experiments are taken in Section 4. Finally, the conclusion is drawn in Section 5.

2. Quantitative Method of Ship Collision Risk

At present, ship collision risk quantification methods include the collision theory analytical calculation method, ship domain method, fuzzy mathematics method, artificial neural network method, data statistics method, evidence theory method, support vector machine method, etc. Because collision risk index is a fuzzy concept, the fuzzy mathematical method is widely used. In this paper, AIS data is used to calculate ship encounter parameters. On this basis, CRI is obtained by the fuzzy membership degree, so as to obtain the time sequence data set required for ship collision risk prediction.

2.1. Calculation of Ship Encounter Parameters. In the process of ship encounter, DCAP, TCPA and other parameters are usually used for quantitative analysis, so as to quantify collision risk. Suppose that the own ship's geographic coordinate is (x_o, y_o), the speed is V_O, and the course is φ_o; the geographical coordinate of the target ship is (x_t, y_t), the speed is V_t, and the course is φ_t. The geometric relation of the parameters is shown in Figure 1, and the calculation formula of relevant parameters [23] is as follows: where D is the relative distance (unit: nautical mile), V_r is the relative ship speed (unit: knot), φ_r is the relative heading (unit: degree), and q is the relative bearing (unit: degree). The value of the arc-tangent function needs to be based on the position relationship, and the range is $[0, 360°]$. In Reference [24], DCPA is the minimum encounter distance (unit: nautical mile), and TCPA is the minimum encounter time (unit: minute). Considering the calculation requirements of the membership function, DCPA and TCPA are greater than zero in this paper, and absolute values are taken when there are negative values.

$$D = \sqrt{(x_t - x_o)^2 + (y_t - y_o)^2},$$

$$V_r = V_O \sqrt{1 + \left(\frac{V_t}{V_O}\right)^2 - \frac{2V_t}{V_O}\cos(\varphi_t - \varphi_o)},$$

$$\phi_r = \cos^{-1}\left(\frac{V_O - V_t\cos(\varphi_t - \varphi_o)}{V_r}\right),$$

$$q = \arctan\left(\frac{x_t - x_o}{y_t - y_o}\right),$$

$$\text{DCPA} = D.\sin(\phi_r - q - \pi),$$

$$\text{TCPA} = D.\cos\frac{(\phi_r - q - \pi)}{V_r},$$

$$\tag{1}$$

DCPA and TCPA are the main parameters used to measure collision risk. When DCPA and TCPA are less than the safety value, there may be a collision. However, to obtain the collision risk, it is not enough to consider only DCPA and TCPA. Other factors, such as distance and relative orientation, etc., should also be considered.

2.2. CRI and Membership Function. CRI is used to measure ship collision risk. Usually, the value range can be $[0, 1]$. The higher the value, the more dangerous it is. DCPA, TCPA, D, q and k (ratio of the speeds of the other and own ship) can be used as the constituent index to measure collision risk, which is obtained by weighting the membership degree in fuzzy mathematics. This paper refers to the membership
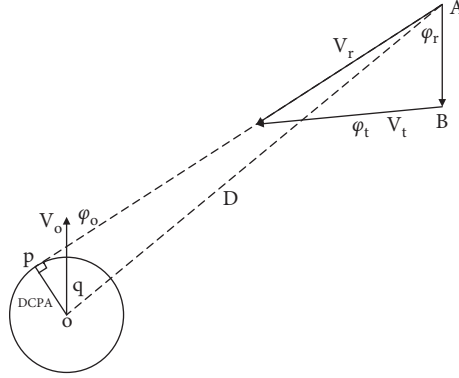
FIGURE 1: Geometric relation diagram of ship encounter parameters.

function in Reference [13] and the weighting coefficient which is the survey results of the experts in Reference [25].

The membership function and weighting coefficient of the above parameters are as follows:

$$
\mu(\text{DCPA}) = \begin{cases} 1, & \text{DCPA} \leq d_1, \\ \dfrac{1}{2} - \dfrac{1}{2}\sin\left[\dfrac{\pi}{d_2 - d_1}\left(\text{DCPA} - \dfrac{d_1 + d_2}{2}\right)\right], & d_1 < \text{DCPA} < d_2, \\ 0, & d_2 \leq \text{DCPA}, \end{cases} \tag{2}
$$

$$
\mu(\text{TCPA}) = \begin{cases} 1, & \text{TCPA} \leq t_1, \\ \left(\dfrac{t_2 - \text{TCPA}}{t_2 - t_1}\right)^2, & t_1 < \text{TCPA} \leq t_2, \\ 0, & t_2 < \text{TCPA}, \end{cases} \tag{3}
$$

$$
\mu(D) = \begin{cases} 1, & 0 < D \leq D_1, \\ \left(\dfrac{D_2 - D}{D_2 - D_1}\right)^2, & D_1 < D \leq D_2, \\ 0, & D_2 < D, \end{cases} \tag{4}
$$

$$
\mu(q) = \dfrac{17}{44}\left[\cos(q - 19^\circ) + \sqrt{\cos^2(q - 19^\circ) + \dfrac{440}{289}}\,\right], \tag{5}
$$

$$
\mu(k) = \dfrac{1}{1 + (k_0/k)^2}, \tag{6}
$$

where d_1 is the minimal safe encounter distance, and d_2 is the absolute safe encounter distance in (2). t_1 and t_2 are the time to reach the corresponding positions of D_1 and D_2 by relative motion in (3). D_1 is the latest steering distance, and D_2 is the moving boundary in (4). The above parameters are calculated according to Reference [12] and Reference [23]. In (5), when q is equal to 19°; the collision danger degree is the highest. k_0 is usually taken as one in (6).

Weighting the membership degree of the above parameters, the collision risk index formula is as follows:

$$
\text{CRI} = \mu(\text{DCPA}) * w_1 + \mu(\text{TCPA}) * w_2 + \mu(D) * w_3 + \mu(q) * w_4 + \mu(K) * w_5, \tag{7}
$$

where the values of weight coefficients $w_1, w_2, w_3, w_4,$ and w_5 are 0.36, 0.32, 0.14, 0.10, and 0.08, respectively.

3. Collision Risk Prediction Model

3.1. LSTM Model. LSTM neural network, namely long and short memory neural network, is an improved network of RNN [26]. According to Reference [27], LSTM model can remember past information and store it for a long time by adding additional memory units, which has strong generalization ability, good learning ability for both large and small data sets, and strong advantages in dealing with nonlinear problems. LSTM model can reflect the time dependence and correlation in time sequence data of AIS date and predict ship collision risk. The main structure of LSTM model has three gates. Its basic structure is shown in Figure 2.

The forget gate f_t determines the information to be discarded and retained according to the unit state C_{t-1} at the previous moment, and the input x_t determines the value to be updated through σ and tanh respectively and generates new candidate values for updating [27]. The value after the updating operation will be updated together with the Forget Gate f_t, and the updated unit state C_t is computed with the tanh function and Output Gate o_t and then outputs h_t [28]. The state update equation of LSTM basic unit are equations (8)–(13).

$$f_t = \sigma\left(W_f.[h_{t-1}, x_t]\right) + b_f, \tag{8}$$

$$i_t = \sigma\left(W_f.[h_{t-1}, x_t]\right) + b_i, \tag{9}$$

$$a_t = \tanh\left(W_c.[h_{t-1}, x_t]\right) + b_c, \tag{10}$$

$$C_t = f_t.C_{t-1} + i_t.a_t, \tag{11}$$

$$o_t = \sigma\left(W_o.[h_{t-1}, x_t]\right) + b_o, \tag{12}$$

$$h_t = o_t.\tanh\left(C_t\right), \tag{13}$$

where x_t is the input at time t; f_t, i_t, o_t represent forget, input and output gate respectively. a_t is the input node state at the corresponding moment; C_{t-1} and C_t are the unit states at the corresponding moment. h_{t-1} and h_t are the outputs at corresponding moment. σ is the sigmoid activation function, and tanh is a hyperbolic tangent function [28]; $W_f, W_i, W_c,$ and W_o and $b_f, b_i, b_c,$ and b_o are the corresponding weight matrices and offsetting vectors.

3.2. PSO Algorithm. PSO algorithm is an optimization technology proposed according to the foraging behavior of birds, which updates the formula with speed to make the particles in the population constantly close to the historical optimal value [29]. Firstly, the particle is given an initial velocity and position information in the solution space in a random way, and then the local and global optimal solutions are tracked and their velocities and positions are updated in time through the defined Fitness function. In this process,
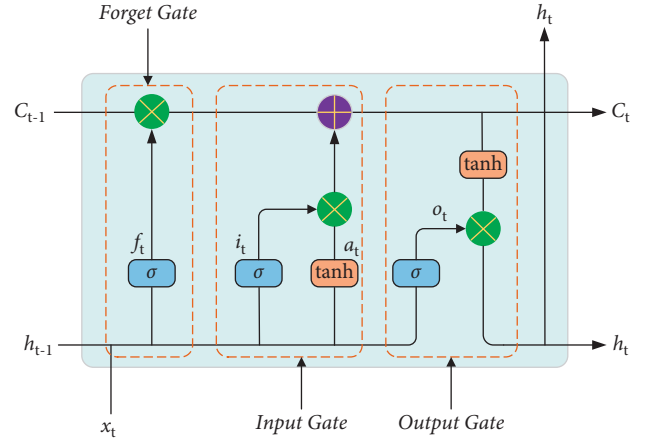
FIGURE 2: Basic structural unit of LSTM.

the Fitness value is calculated through each iteration, so as to achieve the set target Fitness value, and then the optimal solution can be obtained.

In the search space, several particles form a population, and after the t iteration, the velocity and position of particles are formed, which are represented by $X_{i,t}$ and $V_{i,t}$ respectively. During the search process, the position and velocity are updated constantly, and two optimal solutions are formed: one is the individual extremum $pbest_i$, and the other is the global optimal solution $gbest_i$. In the process of searching for the optimal solution, the particle updating formulas of velocity and position [30] are (14) and (15), respectively.

$$\begin{aligned} X_{i,t+1} = {}&\omega X_{i,t} + c_1 \text{rand}\left(pbest - X_{i,t}\right) \\ &+ c_2 \text{rand}\left(gbest - X_{i,t}\right), \end{aligned} \tag{14}$$

$$X_{i,t+1} = X_{i,t} + \lambda V_{i,t+1}, \tag{15}$$

where ω is the inertia factor; c_1 and c_2 are individual and global learning factors, respectively. *rand* is a random number in $[0, 1]$. λ is the velocity coefficient.

3.3. PSO-LSTM Model for Collision Risk Prediction. As hyperparameters greatly impact on the results, PSO algorithm can improve the accuracy of LSTM model by optimizing parameters [31]. The flow chart of PSO-LSTM prediction model is shown in Figure 3. The steps are as follows:

Step 1: According to the AIS data, the ship motion parameters are obtained and the ship encounter parameters are calculated.

Step 2: CRI is obtained by calculating the membership function value and weight coefficient by fuzzy mathematics, and the training and test data are constructed according to the time sequence. Since the range of CRI is already $[0, 1]$, normalization is not required.
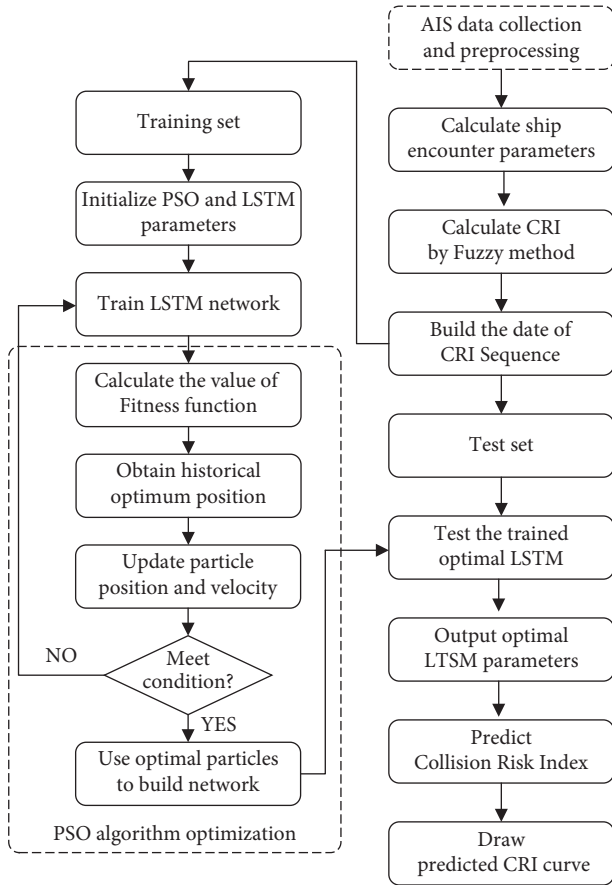
Figure 3: Flow chart of collision risk prediction by PSO-LSTM model.

Step 3: The particle swarm parameters are initialized, including the population size, maximum number of iterations, learning factor, particle positions and velocity ranges [32]. Meanwhile, the hyperparameters of LSTM neural network are initialized, including the number of hidden layer neurons, the learning rate, the maximum number of iterations of LSTM network, and the number of steps in the input layer.

Step 4: The LSTM model is established, which is trained with the data of the training set. The results are compared with the training set, and RMSE is used as the Fitness function in PSO algorithm. The calculation formula is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\hat{Y}_i - Y_i \right)^2}, \tag{16}$$

where Y_i and $\hat{Y}_i$ are the real value of the CRI time sequence and the corresponding predicted value of the model respectively.

Step 5: The global optimal position *gbest* and local optimal position *pbest* are determined by the initial Fitness value of the particles [33], and they are regarded as the historical optimal positions. According to the speed and position of the updated particles, the corresponding particle Fitness

value is calculated and then updated to improve the prediction accuracy.

Step 6: The termination condition is judged to be satisfied (the Fitness value of particles tends to be stable or reaches the maximum number of iterations with iteration [33]). If the termination condition is satisfied, optimal parameters are assigned to PSO-LSTM model; otherwise, step (4) is returned.

Step 7: PSO-LSTM prediction model is constructed from the optimal parameter values, which is applied to predict CRI of the test sets. The predicted result curves are drawn, and the ship collision risk is visualized.

4. Experiment

This paper is based on AIS data of pilot's BRM (bridge resource management, which is included in the Mandatory Enforcement Part A of the STCW convention by IMO in 2010) practice test on ship handling simulator, and the actual operation site is shown in Figure 4. The main reasons for choosing the ship maneuvering simulator for the test are as follows: first, different encounter situations are simulated by setting the same environmental parameters. In the case of the same other conditions, interference is eliminated so as to compare the difference of different encounter situations. Second, the AIS data collected by the simulator have a complete structure and high data frequency, which can well solve the problem of a lack of AIS data and time synchronization. All these are conducive to better process analysis. Finally, experienced pilots perform practical operation experiments, which can not only ensure the professionalism of the experiment, but also ask pilots how they feel in the process of ship encounter, which can be verified and compared with the results of the test data analysis. In addition, the disadvantages of the simulator test also need to be explained, mainly as follows: the environmental conditions in a certain water area are relatively fixed, and are not as complex as the actual water area. The pilot's sense of urgency for dangerous situations is not as strong as the actual situation, which has a certain influence on the judgment of collision risk.

Through data acquisition and processing, ship encounter parameters and fuzzy collision risk index are calculated, PSO-LSTM network is trained to predict CRI, and single-ship encounter situation and multi-ship encounter situation tests are carried out. Considering that disadvantage of PSO-LSTM neural network is that it requires a long period of training, it is necessary to set reasonable parameters to avoid PSO optimization parameter setting too large search range and iteration times, to improve operation efficiency. Therefore, during the experiments, the PSO parameters are set as follows: the learning factors c_1 and c_2 are both set to 2, the range of the inertia factor ω is from 0.8 to 1.2, the maximum number of evolution is 20, and the size of the particle population is 5.

The selected water area for the test is the approach channel of Shanghai Yang Shan Port. The main reasons for

FIGURE 4: Pilot's actual operation in ship handling simulator.

the test in this area are as follows: Shanghai Yang Shan Port is a world-famous container terminal, and its inbound and outbound channels have certain typicality. At the same time, due to the large scale of navigable ships, the navigation safety of their waters is also a concern. The parameters of the own ship and target ships are shown in Table 1. The environmental parameters are shown in Table 2, which need to be marked, because environmental conditions, especially strong winds and waves, will affect the navigation difficulty of ships, affect the pilots' judgment of collision risk, and then affect the generation and prediction results of training datasets.

4.1. Single-Ship Encounter Experiment. AIS data are collected with an interval time of 6 seconds during the encounter between ship OS01 and ship TS05. In the experiment, approximately 255 groups of data is collected and processed, with the first 175 groups of data as training sets and the others as test sets. A complete group of data includes AIS data of the two ships, encounter parameters calculated by the formula, and CRI calculated by the fuzzy method. Taking 10 groups of data as time intervals, corresponding to one minute, are used in the rolling window. After the training and testing PSO-LSTM network, the experimental result curves of DCPA, TCPA and CRI and Fitness, etc., are smoothed and the relevant curves are drawn as follows:

Analysis of the single-ship experiment results is as follows:

(1) Figure 5 reflects the track of ship OS01 and ship TS05 in the encounter process. The arrow represents the direction of the track, and the colour filled in track dots of the target ship TS05 reflects CRI which is obtained according to the Fuzzy calculation method. It can directly reflect the change in collision risk in the encounter process. The redder the colour is, the more dangerous it is.

(2) From Figure 5, the risk index of the target ship leaving the port turns red many times during the encounter with its own ship, indicating that the collision risk of the target ship increases significantly during the close encounter. At this time, the ship's pilot needs to be vigilant and pay close attention to the dynamics of the target ship in the encounter until it passes clear.

(3) According to Figures 6 and 7, in the process of ship encounter, DCPA and TCPA decrease and then increase rapidly as the two ships get closer and farther away. The corresponding collision risk also increases and then decreases.

(4) The Fitness curve in the training process of PSO-LSTM network is reflected in Figure 8, which is also the RMSE convergence process. The experiment shows that the deviation is rapidly reduced and stable in the iterative optimization process of PSO, and PSO-LSTM network training process is effective.

(5) The deviation of PSO-LSTM test prediction results is reflected in Figure 9, which indicates that the CRI prediction accuracy of PSO-LSTM is high, and the overall deviation range is ±0.1.

(6) Figure 10 intuitively reflects the prediction effect of PSO-LSTM model by comparing the original data, the predicted data of training process and the predicted data of testing process. The experimental results show that PSO-LSTM model has strong prediction ability and good prediction effect. It also shows that PSO-LSTM model can be well adapted to ship collision prediction in single-ship encounter situation.

(7) Figure 11 shows the visualization of CRI prediction results. The changes of CRI in ship encounter can be intuitively seen in relative coordinates, where the arrow is the velocity vector of the target ship TS05. From the figure, it can be easily found that when ship TS05 approaches ship OS01, its CRI increases rapidly, and the colour of the position point of ship TS05 turns red rapidly.

(8) After the experiment, the pilots are consulted about their feelings. In the process of meeting with the target ship TS05, the pressure is obviously felt when approaching, which is due to the close distance and the large ship size. The psychological danger perception of the driver at that time is consistent with the situation reflected by the CRI curve in this test.

4.2. Multi-Ship Encounter Situation Experiment. References for the single ship test, the same test parameters are set except for the target ship. AIS data are collected with an interval time of 6 second during the encounter between

TABLE 1: Main parameters of the ships in experiments.

AIS number	Attribute	Length (m)	Width (m)	Draught (m)	Ship type	Encounter situation
OS01	own ship	274.7	40	10	5000TEU container ship	—
TS05	Target ship	347	42.8	14.4	8000TEU container ship	Head-on
TS08	Target ship	81.8	13.8	4.4	3000 ton river ship	Starboard crossing
TS19	Target ship	263.2	32.2	8.5	3000TEU container ship	Portside crossing

TABLE 2: Environmental parameters during the simulation test.

Wind direction (degree)	Wind scale	Flow direction (degree)	Flow velocity
45	Level 4	310	1 knot
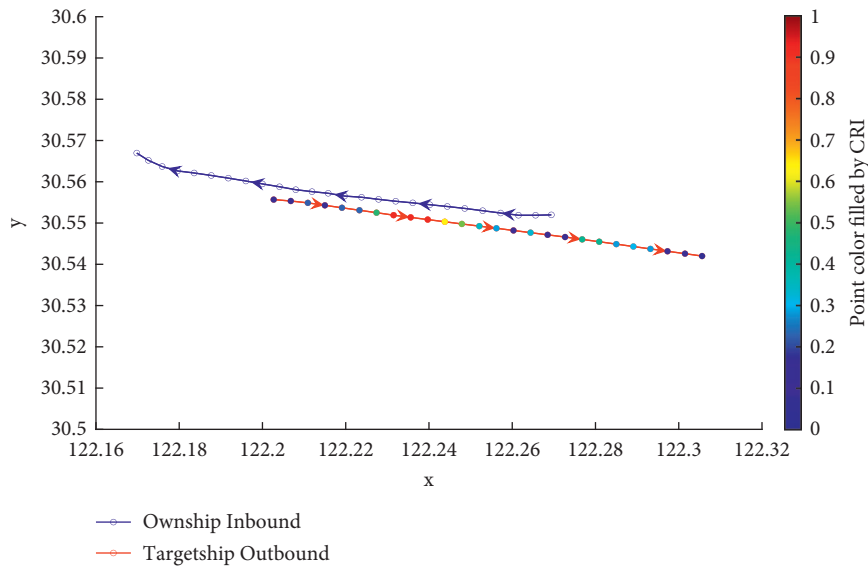
FIGURE 5: Visualization of collision risk between ship OS01 and ship TS05 with head-on situation.
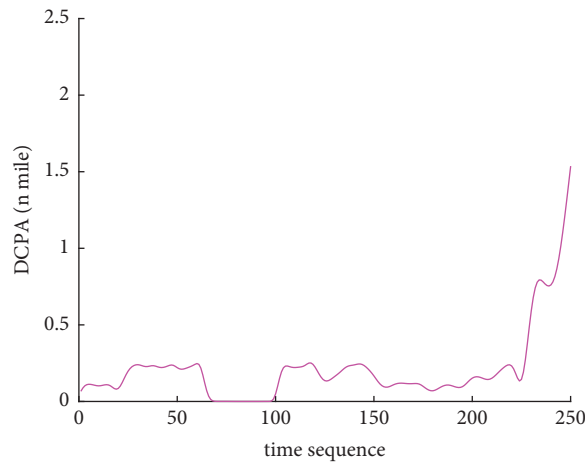
FIGURE 6: DCPA curve of single-ship encounter process.

ship OS01 and ships TS08 and TS19. In the experiment, approximately 255 groups of data is collected and processed, with the first 175 groups of data as training sets and the others as test sets. Taking 10 groups of data as time intervals, corresponding to one minute, are used in the rolling window. After the training and testing PSO-LSTM network, the
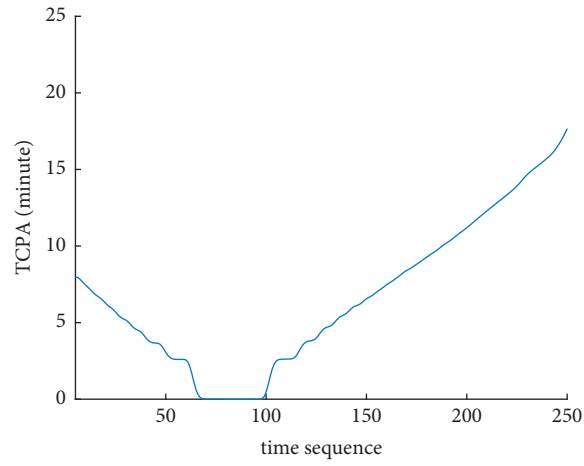
FIGURE 7: TCPA curve of single-ship encounter process.
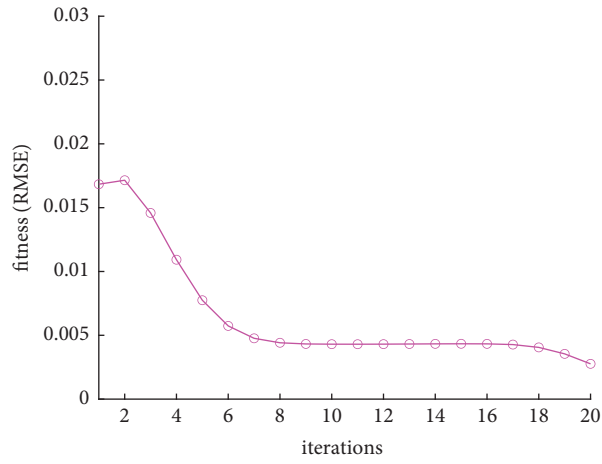
FIGURE 8: Fitness curve of PSO-LSTM training process.
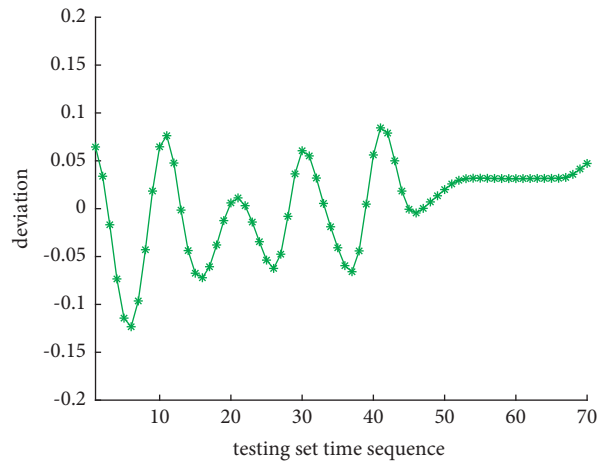
FIGURE 9: CRI deviation curve of PSO-LSTM prediction.
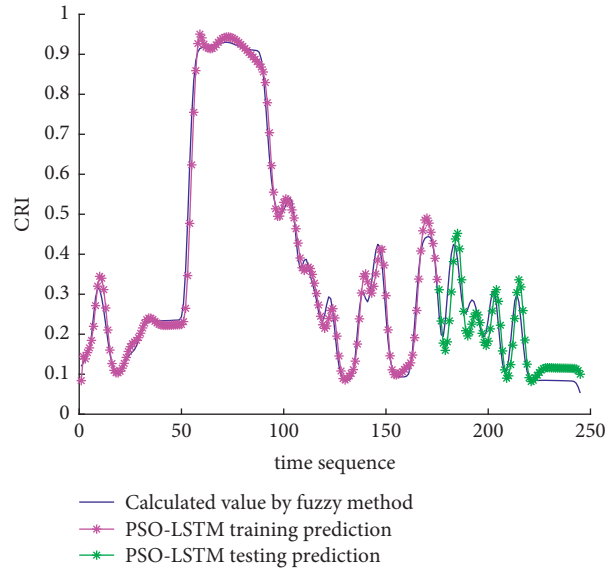
Figure 10: Predictive results of CRI in training and testing.
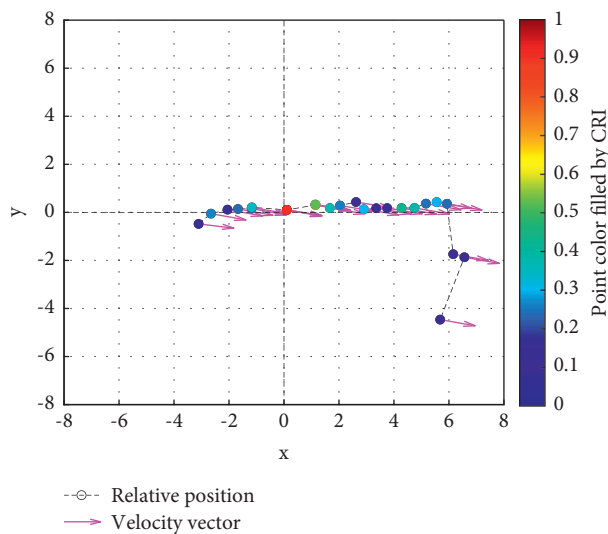
Figure 11: Visualization of collision risk prediction of ship TS05 in relative coordinates of own ship center.

experimental result curves of DCPA, TCPA, CRI and Fitness, etc., are smoothed and the relevant curves are drawn as follows:

The analysis of the multi-ship experiment results is as follows:

(1) Figure 12 reflects the track of the ship and the target ships TS08 and TS19 in the encounter process. According to the colour changes of the position points, the CRI of the target ship TS08 is relatively stable, and the risk increases slightly when the target ship crosses and approaches the own ship from the starboard side, indicating that the collision risk is small. By comparison, the CRI of the target ship TS19 changed relatively greatly, resulting in several sudden increases. In the process of the target ship crossing and approaching from the port side, the corresponding position points turned red several times, indicating the danger of collision. The pilot needs to pay attention to the target ship TS19 until it passes and clears. In the experiment, the overall collision risk of the target ship from the port side is greater than that of ship from the starboard.

(2) Figures 13 and 14 reflect the change characteristics of the DCPA and TCPA, respectively, in the process of multi-ship encounter. By comparison, it is found that with the approach and departure of the target ship, the TCPA decreases roughly and then increases rapidly, while the DCPA shows the characteristics of fluctuation.

(3) In Figure 15, the Fitness curves of target ships TS08 and TS19, also known as RMSE curves, can converge, indicating that the training process of PSO-LSTM is
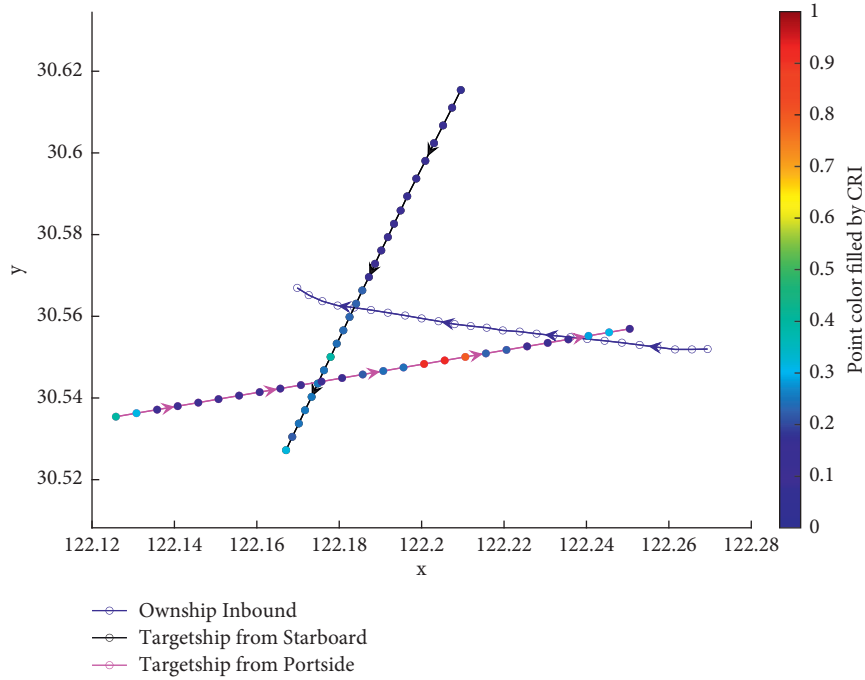
FIGURE 12: Visualization of collision risk in multi-ship encounter situation.
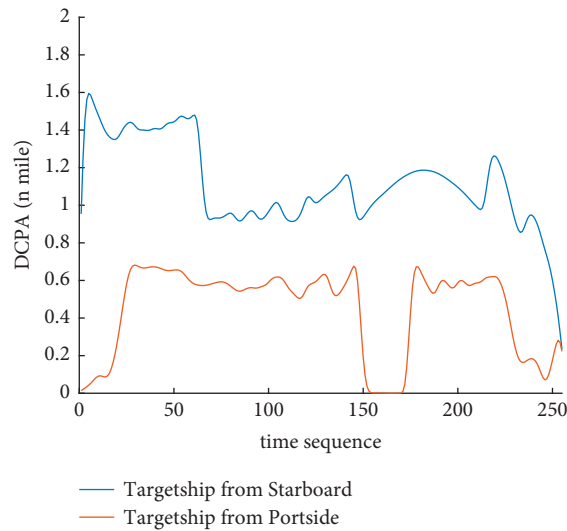
FIGURE 13: DCPA curve of multi-ship encounter process.

effective and that the deviations are gradually reduced under the action of PSO optimization. The Fitness curves show that starboard ship is faster than port ship in PSO optimization and that deviation convergence is faster.

(4) In Figure 16, the CRI prediction curve analysis of the test set shows that the prediction accuracy of PSO-LSTM is high, with the starboard prediction deviation range of ±0.05 and port prediction deviation range of ±0.15.

(5) Figures 17(a) and 17(b) show the performance of PSO-LSTM in the training and testing processes.

There is a high degree of fit between the results and the original data, and the predicted results are close to the actual ones. The results show that PSO-LSTM can effectively and accurately predict ship collision risk, and adapt to the CRI prediction of multi-ship encounter situation.

(6) In Figures 18(a) and 18(b), the CRI change process of target ships TS08 and TS19 is visualized. Figure 18(a) shows that when ship TS08 passed the bow of ship OS01, CRI rapidly and gradually decreased and remained at a low level, indicating a low overall risk. From Figure 18(b), during the encounter between
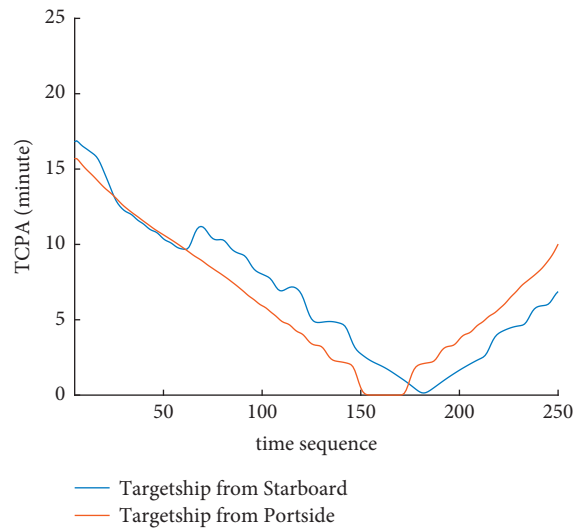
FIGURE 14: TCPA curve of multi-ship encounter process.
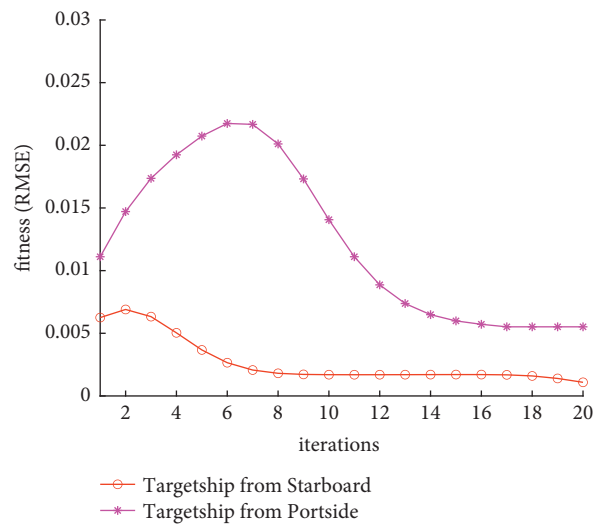
FIGURE 15: Fitness curve of PSO-LSTM training process.
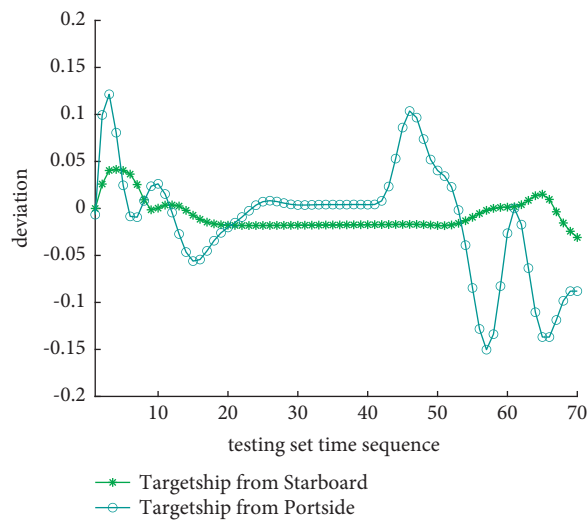
FIGURE 16: CRI deviation curve of PSO-LSTM prediction.
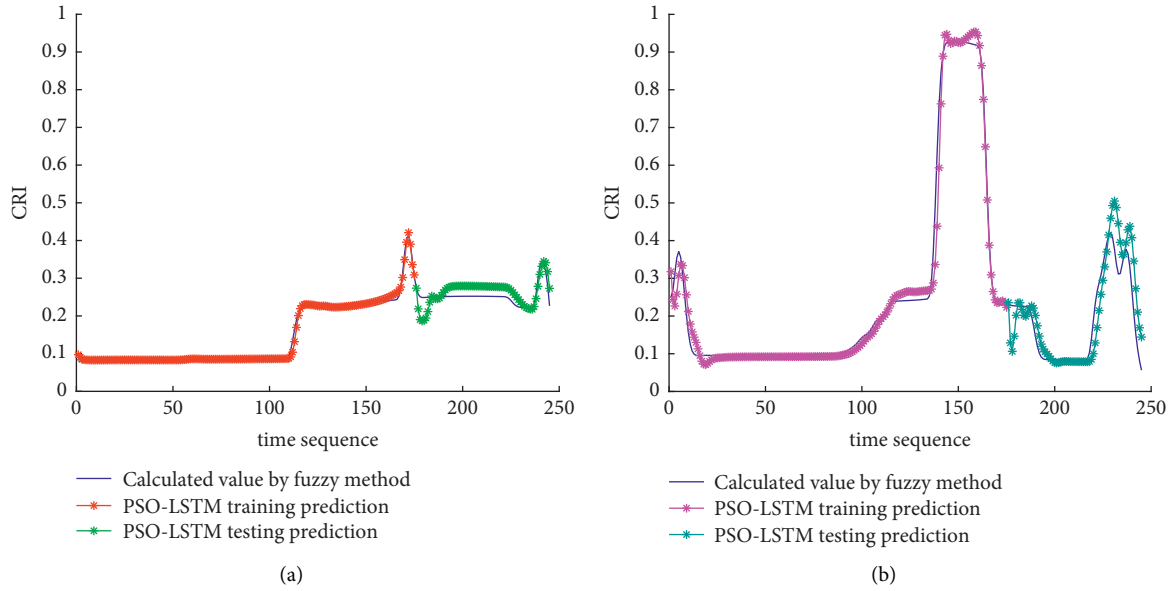
FIGURE 17: Predictive results of CRI in training and testing. (a) Target ship TS08 from Start broad. (b) Target ship TS19 from Portside.
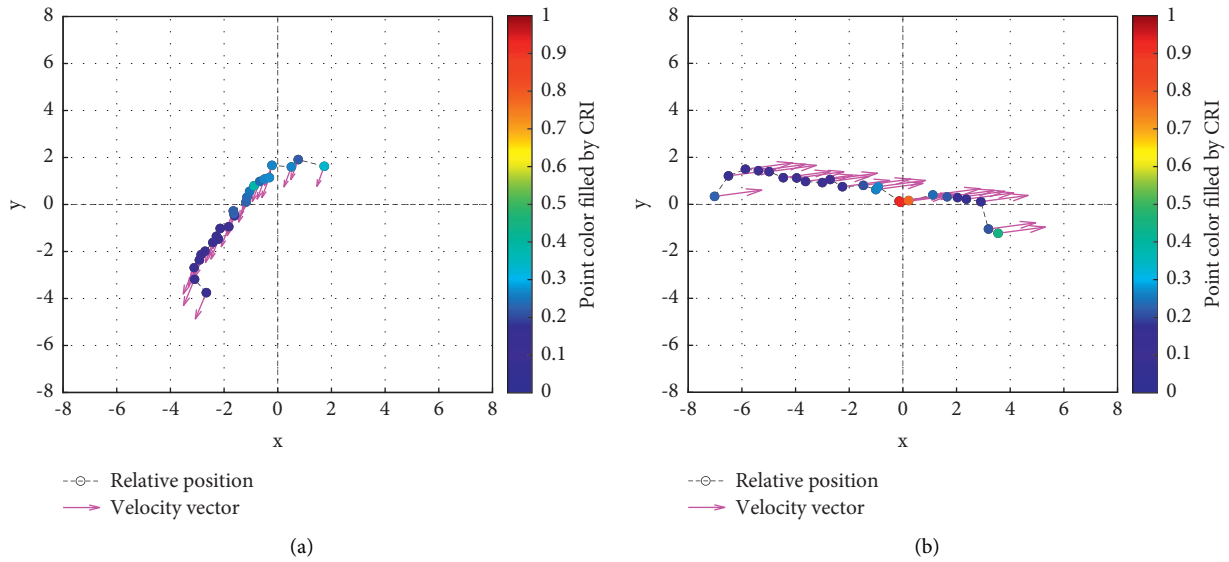
FIGURE 18: Visualization of collision risk prediction in relative coordinates of own ship center. (a) Target ship TS08 from Start broad (b) Target ship TS19 from Portside.

the target ship TS19 and the own ship, there is a very close process, and CRI turns red many times, so there is a collision risk, which should be considered.

(7) After the experiment, the pilots are consulted about their feelings. During the course of the ship encounter, they clearly felt that the ship from the portside was more urgent, closer and more dangerous, which is consistent with CRI and its prediction.

5. Conclusion

Collision risk analysis is an important for ship navigation safety assurance. It is necessary to make collision risk forecast and forewarning. CRI, namely ship collision risk index, is mostly considered with DCPA, TCPA and other parameters in the process of ship encounter. These parameters are time sequences, which in performance before and after time are interrelated, and the corresponding CRI also shows time correlation characteristics. The rolling

prediction mechanism and time memory ability of LSTM model can well reflect the related characteristics. PSO algorithm can optimize the hyperparameters of LTSM model, and better prediction results can be obtained. In this paper, through single-ship encounter and multi-ship encounter tests, PSO-LSTM model was used to obtain good results in CRI prediction. The Fitness curve of PSO in the training process converges rapidly, and the prediction deviation range of the test results is small. The results show that PSO-LSTM model can adapt to complex encounter situations and show high efficiency and accurate ship collision risk prediction.

In the future, we will consider using the multi-ship prediction capability of PSO-LSTM model to access real-time AIS data, and train a more powerful network through massive data learning to provide assistance for collision risk prediction for ship pilots earlier and in a timely manner. In addition, combined with the external environment, the deep learning model will be more effective in predicting and warning ship collision hazards. It is believed that with the development of navigation intelligence in the future, more intelligent deep learning methods will be more widely applied in the navigation field to ensure navigation safety and promote the development of the unmanned vessel and navigation intelligence.

Data Availability

The datasets used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Z. Liu, Y. Li, S. Dong, and Z. Zhang, "Spatial logical relationship analysis model of ship encounter space," *Ocean Engineering*, vol. 239, 2021.

[2] J. Kearon, "Computer program for collision avoidance and track keeping," in *Proceedings of the Conference on Mathematics Aspects of marine Traffic*, pp. 229–242, London, 1977.

[3] L. Kang, Z. Lu, Q. Meng, S. Gao, and F. Wang, "Maritime simulator based determination of minimum dcpa and tcpa in head-on ship-to-ship collision avoidance in confined waters," *TransportMetrica: Transportation Science*, vol. 15, pp. 1–21, 2019.

[4] P. T. Pedersen, "Review and application of ship collision and grounding analysis procedures," *Marine Structures*, vol. 23, no. 3, pp. 241–262, 2010.

[5] R. Zhen, M. Riveiro, and Y. Jin, "A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance," *Ocean Engineering*, vol. 145, pp. 492–501, 2017.

[6] R. Szlapczynski, P. Krata, and J. Szlapczynska, "Ship domain applied to determining distances for collision avoidance manoeuvres in give-way situations," *Ocean Engineering*, vol. 165, pp. 43–54, 2018.

[7] R. Szlapczynski, "A unified measure of collision risk derived from the concept of a ship domain," *Journal of Navigation*, vol. 59, no. 3, pp. 477–490, 2006.

[8] Y. Ren, J. Mou, Q. Yan, and F. Zhang, "Study on Assessing Dynamic Risk of Ship Collision," in *Proceedings of the First International Conference on Transportation Information and Safety (ICTIS)*, pp. 2751–2757, Wuhan, China, 2011.

[9] R. Szlapczynski and J. Szlapczynska, "An analysis of domain-based ship collision risk parameters," *Ocean Engineering*, vol. 126, pp. 47–56, 2016.

[10] Y. Yoo and J.-S. Lee, "Evaluation of ship collision risk assessments using environmental stress and collision risk models," *Ocean Engineering*, vol. 191, Article ID 106527, 2019.

[11] T. Chai, J. Weng, and G. Li, "Estimation of vessel collision frequency in the Yangtze river estuary considering dynamic ship domains," *Journal of Marine Science and Technology*, vol. 25, no. 3, pp. 964–977, 2020.

[12] X. Qu, Q. Meng, and L. Suyi, "Ship collision risk assessment for the Singapore Strait," *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 2030–2036, 2011.

[13] B. Li and F.-W. Pang, "An approach of vessel collision risk assessment based on the D-S evidence theory," *Ocean Engineering*, vol. 74, pp. 16–21, 2013.

[14] J.-H. Ahn, K.-P. Rhee, and Y.-J. You, "A study on the collision avoidance of a ship using neural networks and fuzzy logic," *Applied Ocean Research*, vol. 37, pp. 162–173, 2012.

[15] N. Ho, "Inference rule of collision risk index based on ship near-collision via adaptive neuro fuzzy inference System," *Advances in Science, Technology and Engineering Systems Journal*, vol. 4, no. 4, pp. 152–160, 2019.

[16] K. Zheng, Y. Chen, Y. Jiang, and S. Qiao, "A SVM based ship collision risk assessment algorithm," *Ocean Engineering*, vol. 202, Article ID 107062, 2020.

[17] W. B. Zhang, X. Y. Feng, F. Goerlandt, and Q. Liu, "Towards a Convolutional Neural Network model for classifying regional ship collision risk levels for waterway risk analysis," *Reliability Engineering & System Safety*, vol. 204, Article ID 107127, 2020.

[18] H. C. Chin and A. K. Debnath, "Modeling perceived collision risk in port water navigation," *Safety Science*, vol. 47, no. 10, pp. 1410–1416, 2009.

[19] T. Qin, G. Ma, D. Li, X. Zhou, X. He, and W. Chen, "Dynamic risk prewarning in ship encounter process considering domain violation," *Journal of Navigation*, vol. 74, no. 6, pp. 1416–1431, 2021.

[20] R. Sawada, K. Sato, and T. Majima, "Automatic ship collision avoidance using deep reinforcement learning with LSTM in continuous action spaces," *Journal of Marine Science and Technology*, vol. 26, 2020.

[21] H. Zhou, Yi Zuo, T. Li, and S. Li, "Application of PSO-LSTM for Forecasting of Ship Traffic Flow," in *Proceedings of the 2021 International Conference on Security, Pattern Analysis, and Cybernetics*, pp. 298–302, SPAC, Chengdu, China, 2021, June.

[22] X. Peng and B. Zhang, "Ship motion attitude prediction based on EMD-PSO-LSTM integrated model," *Journal of Chinese Inertial Technology*, vol. 27, no. 4, pp. 421–426, 2019.

[23] T. MingCheng, "Multi-target collision avoidance route planning under an ECDIS Framework," *Ocean Engineering*, vol. 121, pp. 268–278, 2016.

[24] J. Luo, D. Wang, and Y. Gao, "An improved AHP based fuzzy evaluation model for ship collision risk," *Journal of Intelligent and Fuzzy Systems*, 2021.

[25] Z. H. Zhang, *The Research of the Ship Collision Risk Model Based on Fuzzy Comprehensive Evaluation*, Dalian Maritime University, Dalian, China, 2012.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[27] W. Kong, Z. Y. Dong, and Y. Jia, "Etc. Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.

[28] R. Yang, X. Yang, W. Huang, and S. Zhang, *Energy Management of the Power-Split Hybrid Electric City Bus Based on the Stochastic Model Predictive Control*, IEEE Access, USA, 2020.

[29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the . IEEE Int. Conf. Neural Netw*, pp. 1942–1948, Perth, WA, Australia, 1995.

[30] K.-W. Yu and S.-C. Hu, "An Application of AC Servo Motor by Using Particle Swarm Optimization Based Sliding Mode Controller," in *Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics*, Chengdu, China, 2006.

[31] B. Shao, M. Li, Y. Zhao, and G. Bian, "Nickel price forecast based on the LSTM neural network optimized by the improved PSO algorithm," *Mathematical Problems in Engineering*, vol. 2019, Article ID 1934796, 15 pages, 2019.

[32] J. Wang, J. Cao, and S. Yuan, "Shear wave velocity prediction based on adaptive particle swarm optimization optimized recurrent neural network," *Journal of Petroleum Science and Engineering*, vol. 194, 2020.

[33] D. Li, "Human skeleton detection and extraction in dance video based on PSO-enabled LSTM neural network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 2545151, 10 pages, 2021.

Research Article

A Bioinformatics Study of Immune Infiltration-Associated Genes in Sciatica

Tao Ma,[1] Guanhua Li,[1] Yuheng Ma,[1] Zhaoqi Ren,[2] Houyun Xie,[1] Chaoyong Sun,[1] Lei Tian,[1] Hao Zhang,[1] and Wei Wang [1]

[1]*Department of Anesthesiology, PLA Rocket Force Characteristic Medical Center, Beijing, China*
[2]*Department of Infusion, PLA Rocket Force Characteristic Medical Center, Beijing, China*

Correspondence should be addressed to Wei Wang; wangwei_pla@cumt.edu.cn

Sciatica has been widely studied, but the association of sciatica with immune infiltration has not been studied. We aimed to screen key genes and to further investigate the impact of immune infiltration in patients with sciatica. The bioinformatics analyzes were performed based on the GSE150408 dataset. Subsequently, we used CIBERSORT to study the immune infiltration in the disease group. Results showed that 13 genes were with differentially expressions in the sciatica group compared to healthy participants, including 8 up-regulated and 5 down-regulated genes. Through the LASSO model and SVM-RFE analysis, a total of 6 genes have intersections, namely SLED1, CHRNB3, BEGAIN, SPTBN2, HRASLS2, and OSR2. The ROC curve area also confirmed the reliability of this method. CIBERPORT analysis showed that T cell gamma delta infiltration decreased and neutrophil infiltration increased in the disease group. Then the association of these six key genes with immune infiltration was further verified. We found six overlapping genes and found that they were closely associated with the total immune infiltration in the sciatic nerve disease group. These findings may provide new ideas for the diagnosis and therapeutics of patients with sciatica.

1. Introduction

Sciatica is commonly caused by lumbar disc herniation involving peripheral neuropathy [1, 2]. According to statistics, the incidence of sciatica in one's life is as high as 40%. The common treatment methods for sciatica include nonsurgical conservative treatment and nonsurgical treatment. 90% of acute sciatica can be effectively relieved by nonsurgical treatment [3].

Proteomic analysis has identified proteins related to sciatica or intervertebral disc degeneration, which may be involved in the pathophysiological process of sciatica [4]. It is generally believed that mechanical compression combined with immunity and inflammation can lead to sciatica during lumbar disc herniation. Many cytokines related to immunity and inflammation are activated in lumbar disc herniation [5, 6].

In this study, we used two machine learning methods to explore and identify the key genes of patients with sciatica and preliminarily analyzed the immune cell infiltration.

Then further evaluate the correlation between immune cell infiltration and the central gene in sciatica so as to provide new research ideas for the treatment and early detection of sciatica.

2. Materials and Methods

2.1. Screening of Differentially Expressed Genes (DEGs). We downloaded GSE150408 in the GEO database (https://www.ncbi.nlm.nih.gov/geo/). The platform of the GSE150408 mRNA microarray is GPL21185, which was used for the following analyzes.

2.2. Identification of Feature Gene. The feature genes were screened by two machine learning algorithms, least absolute convergence and selection operator (LASSO) and support vector machine-recursive feature elimination (SVM-RFE) and validated in the validation dataset. Machine learning is a new type of algorithm analysis tool. This

study used the method of machine learning to identify features for algorithm analysis. Two machine learning algorithms, LASSO and SVM-RFE, were applied for marker screening. LASSO is a regression algorithm regularized through "*glmnet*" package in R. SVM-RFE is a supervised learning technique that can rank features based on recursion. We adopted the "e1071" package to complete the SVM algorithm.

2.3. Analyzes of Immune Infiltration. The CIBERSORT deconvolution algorithm was adopted for the estimation of different immune cell proportions. Totally, we obtained twenty-two types of immune cells. CIBERSORT filters data with $p < 0.05$. We then calculated each immune cell type's percentage and displayed it as a bar graph. The "pheatmap" package was adopted for the construction of the heat map of the twenty-two types of immune cells. Comparisons of levels of the twenty-two types of immune cells were done using a package.

2.4. Statistical Analysis. Analyzes of the association of immune cells with feature genes were performed using Spearman's rank via *R* software. We used the "ggplot2" package for the visualization of the plot. $P < 0.05$ indicated statistically significant.

3. Result

3.1. Diagnostic Feature Biomarkers Screening. After removing the batch effects, thirteen DEGs were screened out: 8 significantly up-regulated and 5 significantly down-regulated (Figures 1(a) and 1(b)). Using the LASSO regression algorithm, we found 8 potential Figure 1(b) variables for the disease group (Figure 2(a)). A total of thirteen features were determined in Figure 2(b). SLED1, CHRNB3, BEGAIN, SPTBN2, HRASLS, and OSR2 were finally selected in Figure 2(c). Then, ROC was performed for the evaluation of the value of the prediction of the 6 characteristic genes. The AUCs for all 6 genes were greater than 0.8 (Figure 3(a)). It showed that the characteristic biomarkers have a high diagnostic ability, Figure 3(f).

3.2. Analyzes of Immune Infiltration. Immune infiltration in control and sciatica groups was explored with the twenty-two subpopulations of immune cells. The percentage of the twenty-two types of immune cells was visually displayed in Figure 4(a). CIBERPORT analysis showed that T cells gamma delta infiltration decreased and the degree of neutrophils infiltration increased in the sciatica group (Figure 4(b)).

3.3. Relationship of Central Genes with Immune Cells. As shown in Figure 5(a), SLED1 was related to macrophages M0 ($R = 0.38$, $P = 0.025$), B cells memory ($R = 0.42$, $P = 0.014$), neutrophils ($R = 0.56$, $P = 0.00073$) positively and correlated with monocytes ($R = -0.43$, $P = 0.011$), T cells gamma delta ($R = -0.53$, $P = 0.0016$) negatively. CHRNB3 was positively

associated with T cells gamma delta ($R = 0.72$, $P = 3.8e - 06$) and negatively associated with neutrophils ($R = -0.39$, $P = 0.022$), T cells CD4 naive ($R = -0.45$, $P = 0.0081$) (Figure 5(b)).

SPTBN2 had a positive correlation with CD4 memory-activated T cells ($R = 0.38$, $P = 0.025$) and a negative correlation with neutrophils ($R = -0.37$, $P = 0.033$), T cells (CD4 naïve) ($R = -0.41$, $P = 0.017$) (Figure 5(d)). CHRNB3 had a positive correlation with gamma delta-T cells ($R = 0.72$, $P = 3.8e - 06$) and a negative correlation with neutrophils ($R = -0.39$, $P = 0.022$). No correlation was found between BEGAIN, HRASLS2, OSR2, and immune cell infiltration (Figures 5(c)–5(f)).

4. Discussion

So far, there is no specific diagnostic method for sciatica. Combining medical history with a physical examination is the most common diagnostic method [7]. As a common clinical syndrome, sciatica is caused by two causes: one is internal and the other is external factors [8, 9]. When sciatica occurs, it often causes pain in the legs, back and below the knee, usually accompanied by tingling in the legs, numbness or muscle weakness [10, 11].

This study showed that 13 genes were differentially expressed in patients with sciatica. Through two methods, we identified six key genes, which are SLED1, CHRNB3, BEGAIN, SPTBN2, HRASLS2, and OSR2. We determined the association of these differentially expressed genes and immune infiltration in patients with sciatica. CIBERPORT analysis showed that T cell gamma delta infiltration decreased and neutrophil infiltration increased in the sciatica group. Up to now, there has still been a lack of research on brain-enriched guanylate kinase-associated protein. Studies have shown that brain-enriched guanylate kinase-associated protein participates in chronic pain. We demonstrated in the SNI model that mechanical abnormal pain, an abnormal pain condition caused by harmless stimuli, was significantly attenuated in BEGAIN deficient mice [12, 13]. Another key gene found is SPTBN2. It is the research on SPTBN2. At present, it is mainly used in research on congenital cerebellar ataxia and various cancers. In the study of cancer, miR-424-5p was found to be able to accelerate the development of endometrial cancer through regulating SPTBN2 and then the cldn4/PI3K/Akt axis [14–16]. Combined with bioinformatics and cell experiments, SPTBN2 may become a novel target of lung adenocarcinoma. SPTBN2, highly expressed in LUAD, might indicate poor prognosis. Cell experiments confirmed that SPTBN2 could promote the proliferative, migrative, and invasive abilities of LUAD cells [17].

The researchers found that glia was significantly activated in the brains of patients who experienced chronic pain, indicating that immune cells can spread and maintain disease states, including neuropathic pain, through communication with neurons rather than being regarded as bystanders [18]. During nerve injury, neuronal activity will be activated, resulting in the recruitment of monocytes/macrophages (peripheral) to the injured site. At the same
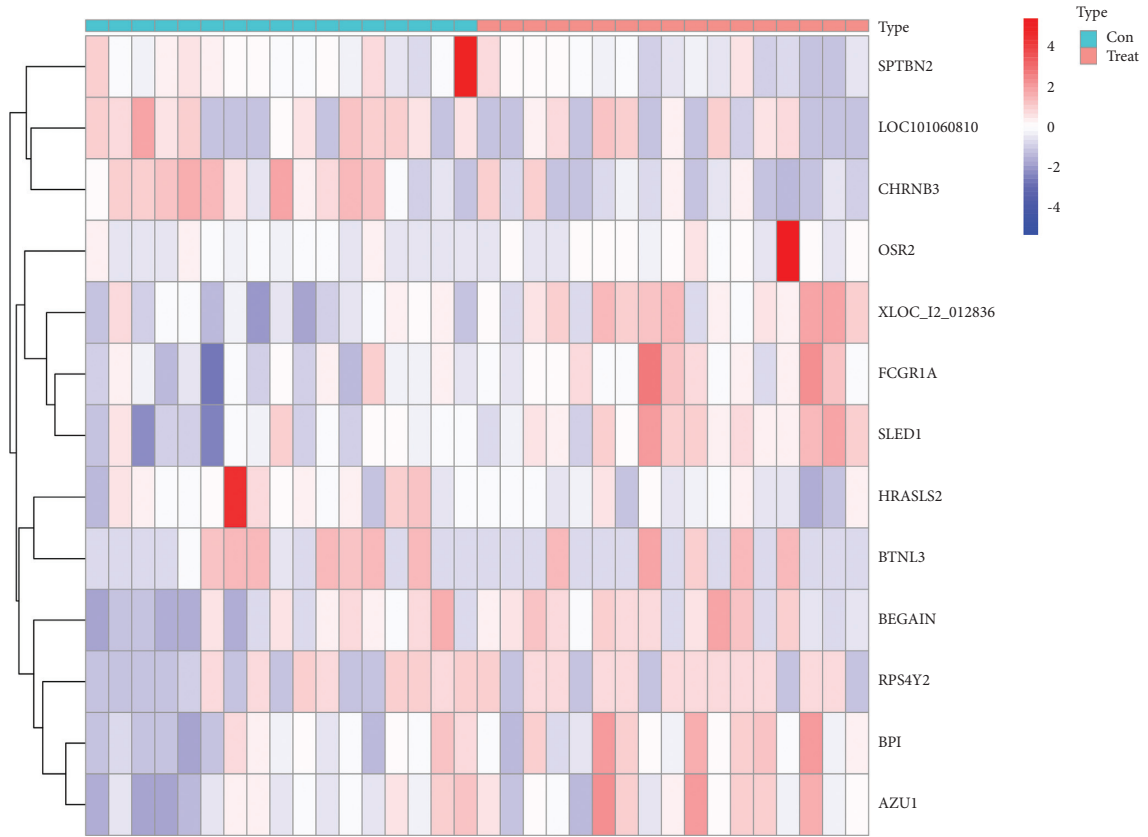
(a)

Figure 1: Continued.

(b)
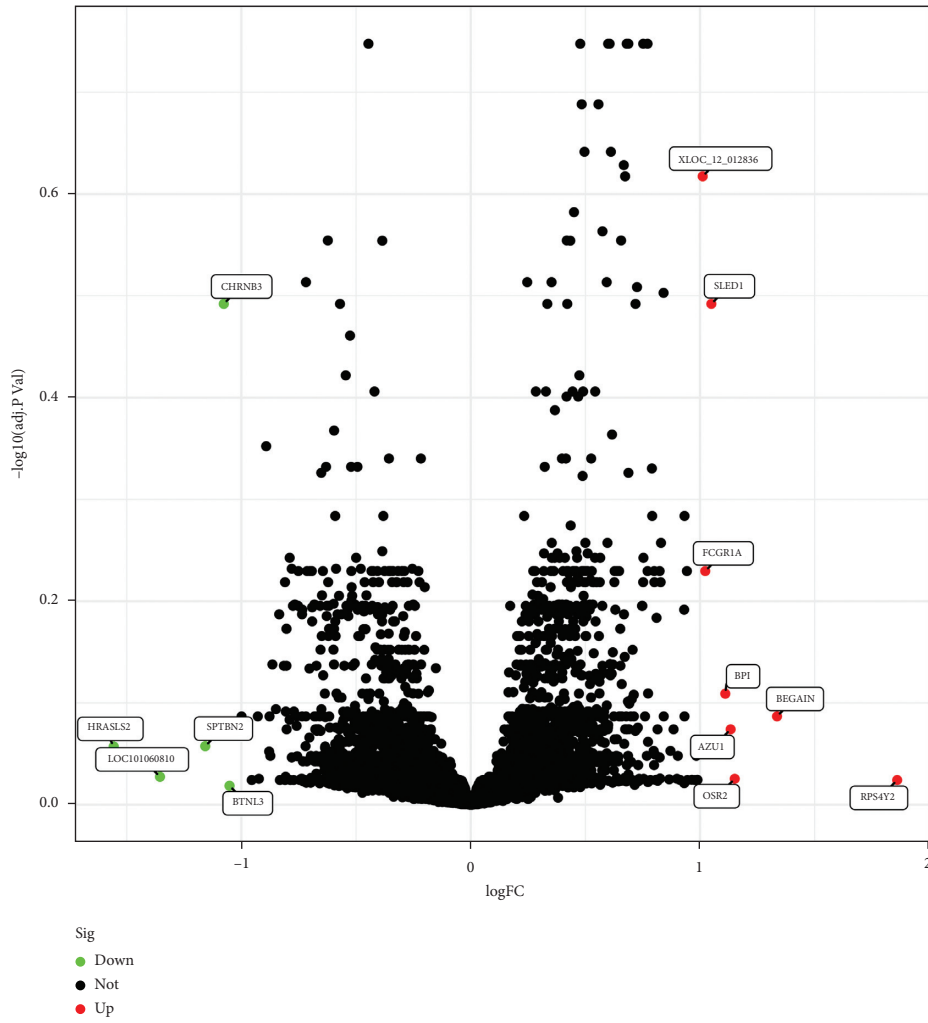
Figure 1: Differential analyzes based on the datasets: (a) heat map of DEGs (adjust $P < 0.05$, |logFC| >1) and (b) the volcano map of DEGs.

(a)   (b)
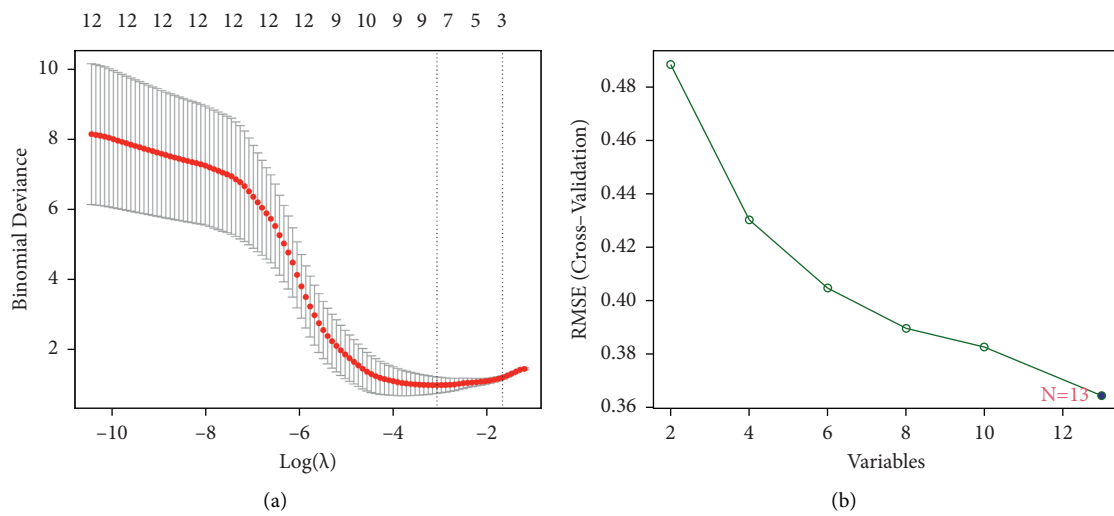
Figure 2: Continued.
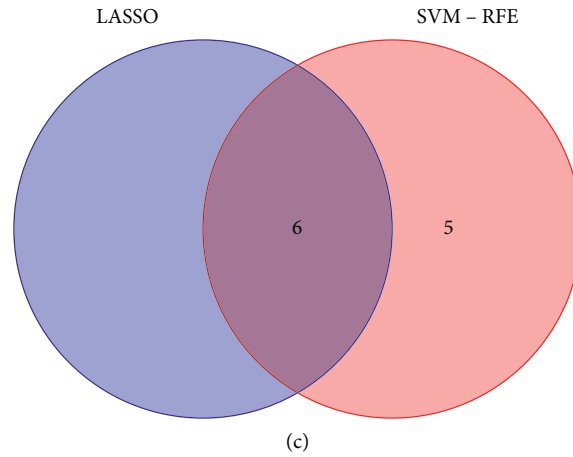
FIGURE 2: Establishment of prognostic genes: (a) selection of tuning feature in LASSO models, (b) selection of biomarkers using SVM-RFE and (c) the overlapped feature genes between 2 methods.
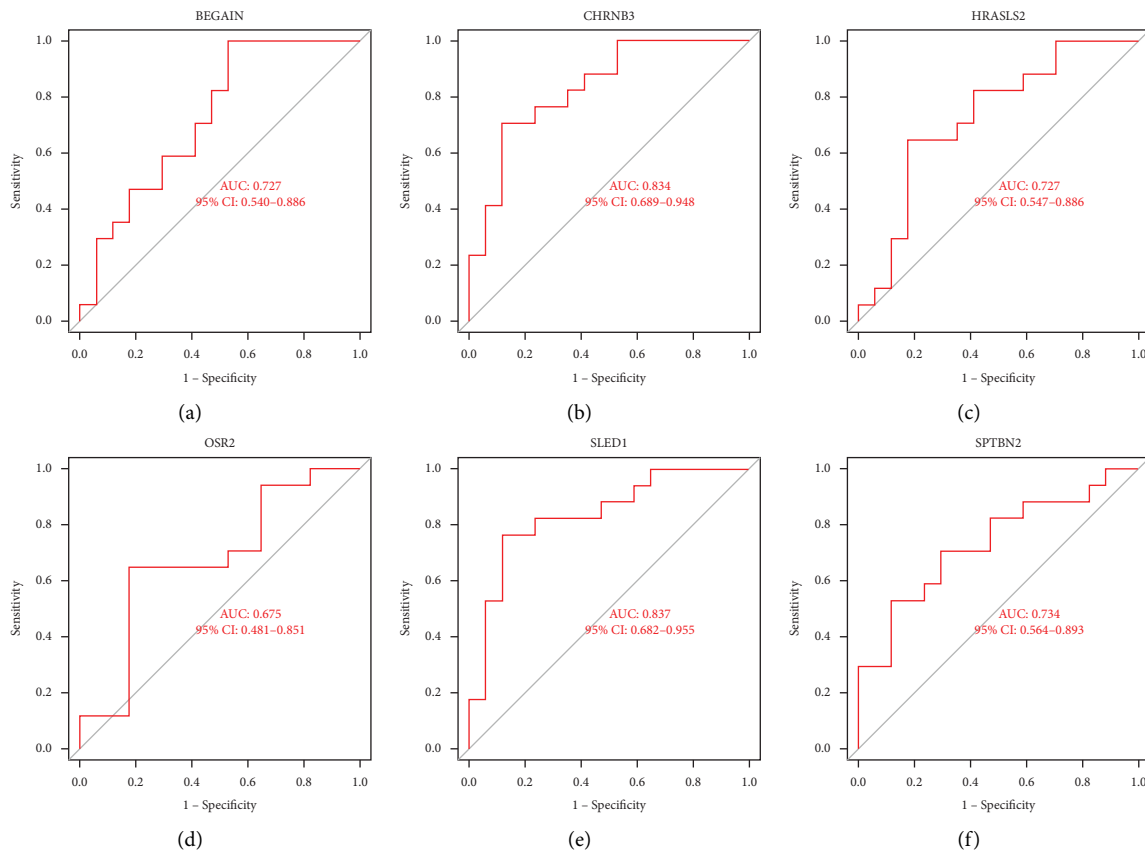
FIGURE 3: Diagnosis efficiency ROC curve of the feature genes: (a)–(f) ROC curve of SLED1 (a), CHRNB3 (b), BEGAIN (c), SPTBN2 (d), HRASLS2 (e), and OSR2 (f).

time, microglia will release inflammatory-related mediators after activation, resulting in neuronal sensitivity [19].

The signal molecules of the immune system are cytokines. An increase of proinflammatory cytokines is related to the existence of pain after nerve injury, while antiinflammatory cytokines are related to the down regulation of the immune system and the relief of neuropathic pain

[20, 21]. Immune system activation has been shown to promote and increase neuropathic pain [22].

Immune cells play an important role in different pathophysiological processes in the state of neuropathic pain. It brings the pain field to different directions and provides opportunities for new methods for the treatment of chronic pain.
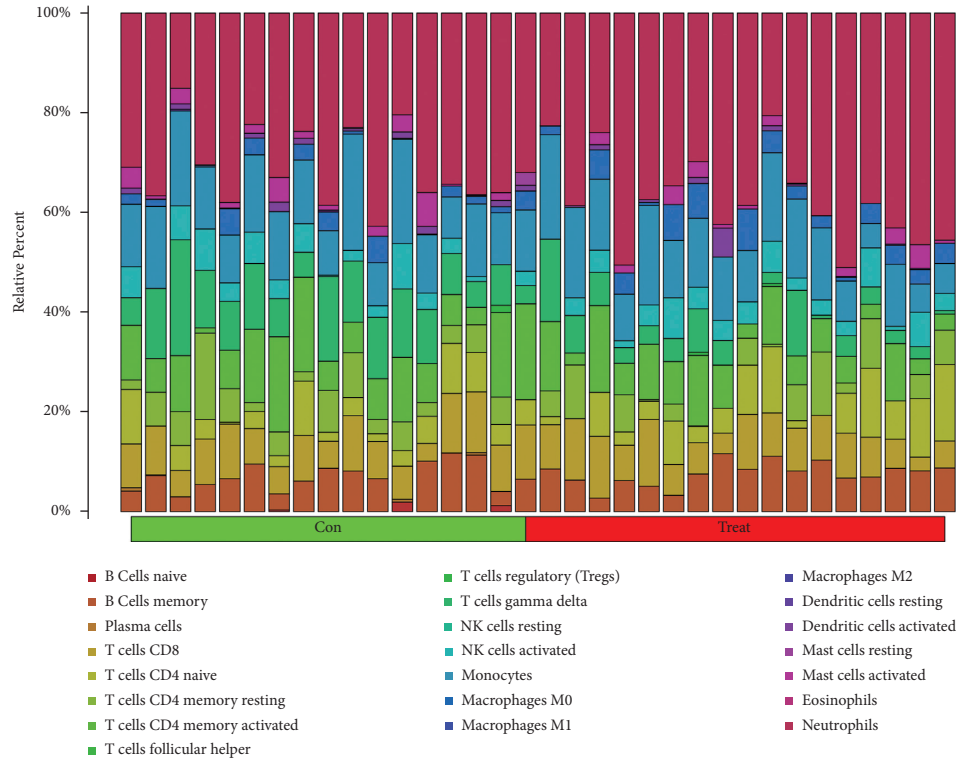
(a)
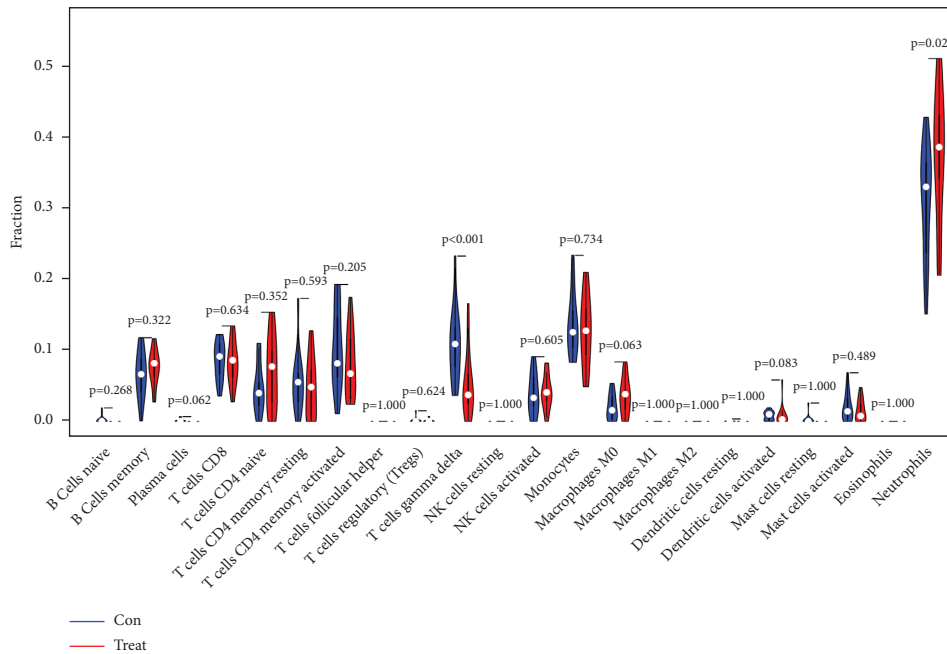
(b)

Figure 4: Visualization and evaluation of immune cell infiltration: (a) relative percentage of 22 kinds of immune cells, and (b) immune cells' violin plot image.

However, there are still some limitations to the present study. This is a purely bioinformatics study without further experiments for validation, which weakened the evidence level of our results. In the future, we will conduct in vivo and in vitro assays to further explore the exact effects of the abovementioned immune-related genes and the potential underlying mechanisms on sciatica.

Figure 5: Correlation analyzes between immune cells and SLED1, CHRNB3, BEGAIN, SPTBN2, HRASLS2, and OSR2. (a) Association of SLED1 with immune cells. (b) Association of CHRNB3 with immune cells. (c) Association of BEGAIN with immune cells. (d) Association of SPTBN2 with immune cells. (e) Association of HRASLS2 with immune cells. (f) Association of OSR2 with immune cells.

5. Conclusions

In summary, we systematically discussed the functions of immune-related genes of sciatica and provided new ideas for new methods for the treatment of chronic pain.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Tao Ma and Guanhua Li contributed equally to this work.

References

[1] M. A. Stafford, P. Peng, and D. A. Hill, "Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management," *British Journal of Anaesthesia*, vol. 99, no. 4, pp. 461–473, 2007.

[2] J.-P. Valat, S. Genevay, M. Marty, S. Rozenberg, and B. Koes, "Sciatica," *Best Practice & Research Clinical Rheumatology*, vol. 24, no. 2, pp. 241–252, 2010.

[3] C. S. Bailey, P. Rasoulinejad, D. Taylor et al., "Surgery versus conservative care for persistent sciatica lasting 4 to 12 months," *New England Journal of Medicine*, vol. 382, no. 12, pp. 1093–1102, 2020.

[4] W. Ren and R. Wu, "Effect of general and sub-arachnoid anesthesia on the incidence of postoperative delirium and cognitive impairments in elderly Chinese patients," *Tropical Journal of Pharmaceutical Research*, vol. 20, no. 2, pp. 433–439, 2021.

[5] R. F. McCarron, M. W. Wimpee, P. G. Hudkins, and G. S. Laros, "The inflammatory effect of nucleus pulposus," *Spine*, vol. 12, no. 8, pp. 760–764, 1987.

[6] K. Hemati, M. Y. Karimi, A. Hosseinzadeh et al., "Induction of analgesia using atorvastatin in experimental diabetic neuropathy through NMDA receptor and inflammatory cytokine inhibition," *Iran Red Crescent Me*, vol. 23, no. 1, p. e454, 2021.

[7] A. H. Ropper and R. D. Zafonte, "Sciatica," *New England Journal of Medicine*, vol. 372, no. 13, pp. 1240–1248, 2015.

[8] B. W. Koes, M. W. van Tulder, and W. C. Peul, "Diagnosis and treatment of sciatica," *BMJ*, vol. 334, no. 7607, pp. 1313–1317, 2007.

[9] I. A. Bernstein, Q. Malik, S. Carville, and S. Ward, "Low back pain and sciatica: summary of NICE guidance," *BMJ*, vol. 356, p. i6748, 2017.

[10] E. R. Stirling, M. S. Patel, and P. J. Sell, "Sciatica," *British Journal of Hospital Medicine*, vol. 77, no. 11, pp. C180–C183, 2016.

[11] Z. Sun, Z. H. Liu, Y. f. Chen et al., "Molecular immunotherapy might shed a light on the treatment strategies for disc degeneration and herniation," *Medical Hypotheses*, vol. 81, no. 3, pp. 477–480, 2013.

[12] T. Katano, M. Fukuda, H. Furue et al., "Involvement of brain-enriched guanylate kinase-associated protein (BEGAIN) in chronic pain after peripheral nerve injury," *eNeuro*, vol. 3, no. 5, 2016.

[13] M. Deguchi, Y. Hata, M. Takeuchi et al., "BEGAIN (brain-enriched guanylate kinase-associated protein), a novel neuronal PSD-95/SAP90-binding protein," *Journal of Biological Chemistry*, vol. 273, no. 41, pp. 26269–26272, 1998.

[14] F. Nicita, M. Nardella, E. Bellacchio et al., "Heterozygous missense variants of SPTBN2 are a frequent cause of congenital cerebellar ataxia," *Clinical Genetics*, vol. 96, no. 2, pp. 169–175, 2019.

[15] M. A. Al-Muhaizea, F. AlMutairi, R. Almass et al., "A novel homozygous mutation in SPTBN2 leads to spinocerebellar ataxia in a consanguineous family: report of a new infantile-onset case and Brief review of the literature," *The Cerebellum*, vol. 17, no. 3, pp. 276–285, 2018.

[16] P. Wang, T. Liu, Z. Zhao, Z. Wang, S. Liu, and X. Yang, "SPTBN2 regulated by miR-424-5p promotes endometrial cancer progression via CLDN4/PI3K/AKT axis," *Cell Death Discovery*, vol. 7, no. 1, p. 382, 2021.

[17] C. Wu, B. Dong, L. Huang et al., "SPTBN2, a new biomarker of lung adenocarcinoma," *Frontiers in Oncology*, vol. 11, Article ID 754290, 2021.

[18] M. Malcangio, "Role of the immune system in neuropathic pain," *Scandinavian Journal of Pain*, vol. 20, no. 1, pp. 33–37, 2019.

[19] H. Tozaki-Saitoh and M. Tsuda, "Microglia-neuron interactions in the models of neuropathic pain," *Biochemical Pharmacology*, vol. 169, Article ID 113614, 2019.

[20] M. A. Thacker, A. K. Clark, F. Marchand, and S. B. McMahon, "Pathophysiology of peripheral neuropathic pain: immune cells and molecules," *Anesthesia & Analgesia*, vol. 105, no. 3, pp. 838–847, 2007.

[21] A. K. Clark, E. A. Old, and M. Malcangio, "Neuropathic pain and cytokines: current perspectives," *Journal of Pain Research*, vol. 6, pp. 803–814, 2013.

[22] K. Ren and R. Dubner, "Interactions between the immune and nervous systems in pain," *Nature Medicine*, vol. 16, no. 11, pp. 1267–1276, 2010.

Retraction

Retracted: The Construction of Big Data Computational Intelligence System for E-Government in Cloud Computing Environment and Its Development Impact

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Ma, F. Hao, Y. Lin, and Y. Liang, "The Construction of Big Data Computational Intelligence System for E-Government in Cloud Computing Environment and Its Development Impact," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7295060, 9 pages, 2022.

Research Article

The Construction of Big Data Computational Intelligence System for E-Government in Cloud Computing Environment and Its Development Impact

Shengqing Ma [1], **Fei Hao**,[2] **Youwu Lin**,[3] **and Yinjing Liang**[4]

[1]*School of Economics and Management, Qinghai Normal University, Xining, Qinghai 810000, China*
[2]*School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China*
[3]*School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China*
[4]*College of Marxism, Taiyuan University of Technology, Taiyuan, Shanxi 030600, China*

Correspondence should be addressed to Shengqing Ma; msq@qhnu.edu.cn

The traditional E-government big data system fills and classifies algorithms with low accuracy and poor work efficiency. With the development and wide application of big data, the internet of things, and other technologies, the integration of information resources has become the key to information construction. In the process of information resource integration, there are still outstanding problems such as incomplete government information resource system, different standards of government information resource management system construction, and serious threats to network and information security. In order to solve this problem, a new E-government big data system filling and classification algorithm is studied in the cloud computing environment; E-government big data filling is carried out on the basis of complete compatibility theory; and the E-government big data computing intelligence system in the cloud computing environment is constructed and its development impact, so as to parallelize the data, classify the data through decision trees, and realize incremental update decision forest parallelization processing. To verify the effectiveness of the method, comparative experiments are set, and the results demonstrate that experiment one is randomly built into the classification model, and according to the decision forest algorithm, the optimal number of decision trees is 24.

1. Introduction

With the rapid development and widespread use of big data technology, people are increasingly optimistic about the future of big data [1]. In particular, the increasing volume of data in recent years has made it difficult for existing software to process it effectively, and the quality of big data makes "big data" technology stand out from the crowd. Big Data can help people extract valid information from large amounts of data quickly and help them get answers quickly [2]. McKinsey & Company, the first to coin the phrase "big data era," has pointed out that data is everywhere and that its value today is not just information but also a factor of production. Deeper mining and organization of data can help people obtain data products. "Big data strategy" is also

on the agenda with the widespread use of big data [3]. Big data technology can be used to collect, process, and handle large amounts of data and analyze them from a professional perspective. People can obtain valuable information.

The use of big data technology for the construction of government information resources integration began to involve the construction of the platform, the construction of the information security system, the promotion of big data awareness, and other aspects, not a quick fix [4]. Big datum is more active in the construction of platforms, and many local governments are building local big data platforms to integrate government information resources across platforms and data types through the construction of big data platforms [5]. Big data platforms can help us gather a large amount of data and analyze and evaluate these data [6]. We

can understand the meaning behind these data. If we just look at the data, it is sometimes difficult to get useful information from it. Big data platforms help managers change the way they used to make decisions based on personal experience and personal knowledge, and through the analysis of data, decisions can be justified, and decision-makers can quickly understand the situation through data analysis [7].

Although big data technology is widely used in the integration of government information resources, there is still a lack of big data awareness in the concrete implementation [8]. Although claiming to use big data to solve problems, the original approach is still used in problem-solving. All these show a lack of awareness of big data and the need to fundamentally understand the importance of big data and the position it occupies in the process of government transformation. At the same time, data security has become an inescapable issue as we make extensive use of data today [9, 10]. The nature of big data allows it to respond quickly to government information, which is why it is so effective in advancing the integration of government information resources so quickly.

The amount of data in the data centre is huge; therefore, it is necessary to improve the efficiency of data processing. Therefore, the paper proposes an E-government big data system filling algorithm to address the problem of data loss that occurs when the government uses the E-government big data system, which solves the problem of missing data, and proposes a parallel optimization of E-government big data system filling based on the cloud computing environment, which greatly improves the efficiency of filling missing data [11]. A new classification model for E-government big data systems in a cloud computing environment is constructed in the paper. By using a decision forest algorithm to classify and process huge data and incremental update processing of the decision forest algorithm, the big data system can continuously change its model according to the continuous change of data, and finally, the parallel optimization of the algorithm is carried out, which makes the processing efficiency greatly improved [12]. The results of the experimental demonstration of the proposed method show that the E-government big data system in the cloud computing environment proposed in the paper has great advantages over the algorithms used in the traditional way.

2. Related Work

The application of big data in government information integration will promote efficient use and innovation of government information [13]. In conclusion, big data are data that transcend traditional data analysis and processing capabilities and are difficult to apply to existing database architectures, and traditional software tools are difficult to access, store, manage and analyze, and thus, new tools and processing methods should be adopted. At the same time, big data are changing the traditional sampling data to all data by changing the precise directional hybrid transformations required from the previous precision to hybridity, more

analysis, and speculation about the trend of things rather than the previous focus on cause and effect.

Yan et al. [14] propose that digital information resources have changed in the big data environment and discuss the three directions of cloud-based information resource services and active innovative service models based on relational networks and Web 2.0. Wei [15] takes the concept and characteristics of cloud computing as an entry point and uses a two-layer perspective to build a top-level design model for information technology. Long et al. [16] think in the direction of using multiple subjects to develop government information resources; propose three new paradigms of using information resources, such as the single information flow model of users, the collaborative model, and the two-way information flow model of users; and explains that the collaborative model of developing and using community information resources has obvious advantages. Fu and Sun [17] discuss the concept of government information resource exploitation and development and propose different positions of public interest and commercial exploitation of government information resource development models according to the development and exploitation models of European and American governments, aiming to give suggestions on the problems and current situation of government information resource development and exploitation [18, 19]. Through the analysis of the development of information resources in China, the importance of knowledge management in the management of government information resources is elaborated, and the strategy of using knowledge management in the development and utilization of government information resources is proposed. Through a survey of 300 valid questionnaires on the impact of big data on community management and influence, the majority of people affirmed that big data can help the government improve scientific and rational decision-making [20, 21].

Literature [22] put forward the idea of further refinement and transformation of China's Comprehensive E-government Thesaurus, providing suggestions for further improvement and changes to the Comprehensive E-government Thesaurus. Hernandez-Marin et al. [23] propose the idea of further refinement of China's Comprehensive E-government Thesaurus. Literature [24] explained the two major government metadata standards DC__government and GILS in the world today from a holistic perspective and put forward suggestions for building a government metadata architecture in China. There are also some scholars who have applied theme maps to the integration of government information resources, providing new methods and approaches to the integration of many E-government information resources, such as [25] and others who have applied theme maps to government portals for analysis and research. Through the analysis and consideration of methods and technologies related to E-government information resources, we can easily see that the current research on the integration of government information resources in China is mostly in the theoretical aspect, while there are not many applications for specific government information resources integration research. In the face of the vast amount of information resources, we should study and explore in depth,

using the emerging big data, cloud computing, and internet of things technology to find effective solutions.

3. Government Information Resources Integration Solutions

In the big data environment, big data and cloud computing technology bring new development opportunities to the integration of government information resources. Big data technology can effectively solve the problems of departmental compartmentalization and unequal information standards in the previous E-government era, and big data technology will undoubtedly promote the development of China's government management in the direction of intelligence. Our government has been actively planning the construction of a big data platform for government information. By analyzing the current situation and problems of government information resource integration and construction under the big data environment, we propose complete solutions to the problems existing in government information resource integration [26].

3.1. Principles of Integration of Government Information Resources. As shown in Figure 1, in the process of integration of government information resources, the government should actively play the leading role of policy guidance, organization and collaboration, standardization, and control of the overall situation to promote the further development of government information resources. At the same time, the integration of government information resources should not only meet the requirements of government departments for the integration of information resources but also meet the requirements of the general public for government information. We should put the principle of serving the people into practice, think what the people think, and sense the urgency of the people. Starting with the most urgent need of the people, we should actively prioritize the integration of government affairs, transportation, and medical services, which are the most popular areas, and gradually realize the full implementation and development of the integration of government information resources [27].

In turn, it eliminates the problem of information silos, thus realizing the integration and sharing of information resources across industries, fields, and shutters and enhancing the effectiveness of government information utilization and the scientific nature of decision-making. Government information can hardly play its proper role if it is only within the government, and there is the use of enterprises to fully explore its potential value but also to further stimulate the demand of potential consumers, which in turn promotes the opening and use of information. While promoting the development of information technology, it also leads to the development of new information industries and the creation of new economic growth points, making the interconnection of information more fluid [28].
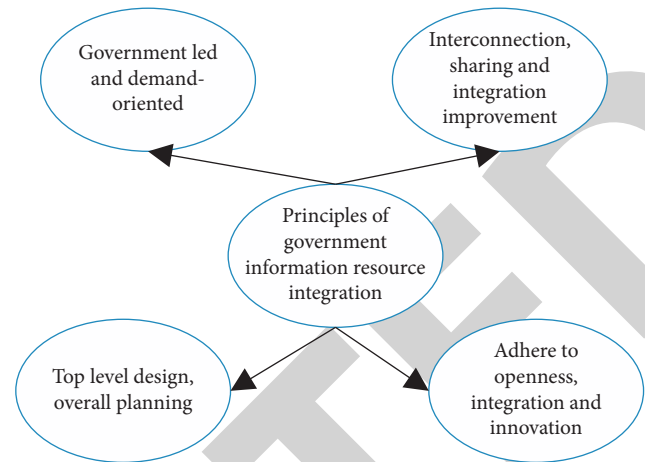
Figure 1: Principles of government information resource integration.

3.2. Government Information Integration Framework and Content Building. The big data platform for government affairs is mainly built by six parts, with the standard specification and information security system as the guarantee and the corresponding platform construction for infrastructure as well as database system and application system as shown in Figure 2.

The public information big data system is based on the collection of information, data exchange, and integration of information to the final information service, with the integration of the following main parts, as shown in Figure 3.

The integration of government information resources is carried out at three main levels, based on standard specifications and secure application systems for infrastructure, systems, and related applications.

4. Big Data System for E-Government in Cloud Computing Environment

In today's cloud computing environment and big data era, the scale of energy consumption big data in the big data centre of E-government is very large; often data collection information is lost during equipment failure, unexpected power failure, and other unstable factors, resulting in the lack of data information and damage, causing losses; the previous way to solve this difficulty usually uses rough set theory. However, a rough set theory can only handle small batches of data, and the processing is very tedious [29]. Based on this, a complete compatibility theory extended from the theory of compatibility relations is proposed in this paper, and based on this theory, an algorithm for filling the E-government big data system is proposed to solve the problem of missing attribute values of data in E-government data centres [24]. The management architecture of the E-government big data system is shown in Figure 4.

As can be seen in Figure 4, the entire management architecture contains a data centre, cluster monitoring module, and sensors. The raw data set is obtained using cluster monitoring to obtain the raw data set that can be used for operation and provide data for later data filing and
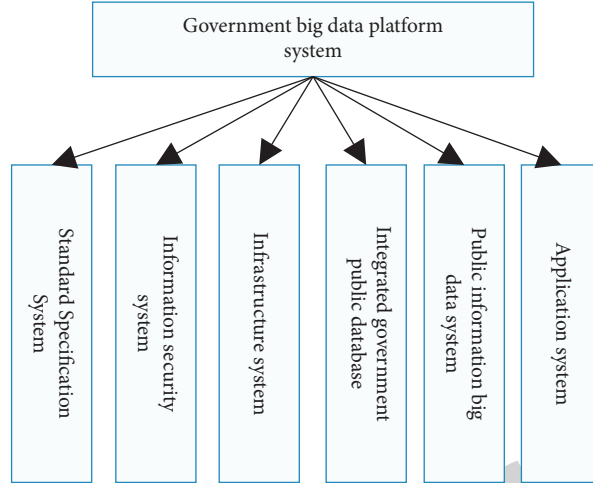
FIGURE 2: Big data platform system for government services.
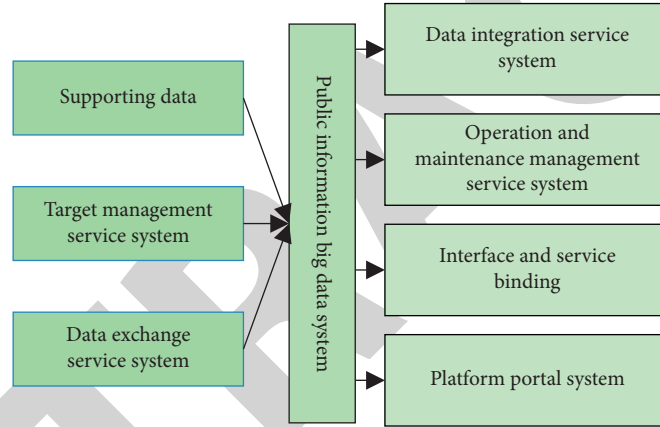
FIGURE 3: Big data system for public information.

classification. The paper proposes an E-government big data filling algorithm based on the theory of completeness and compatibility as a solution to the problem of missing data in E-government data centres. The main process of the process is to discrete the attribute values of the data, interrupt their continuity, screen out the missing data information for separation processing, arrange the attribute values for processing, then perform inverted indexing processing, and then detect whether the missing data are perfect-compatible class data; if it is perfect compatible data, the principle of minimum value is applied, and the corresponding decision attribute generated by the attribute of the corresponding condition is used to fill the missing data. If the missing data are incomplete compatibility data, the attribute value with the highest frequency of the missing attribute is used to fill the missing attribute [25].

In the process of filling in missing data, the data set is decomposed by clustering using the double clustering method, and the data set is divided into multiple data clusters with different characteristics according to the differences in the data characteristics, using the clustering idea that the smaller the average residual within the data cluster,

the higher the similarity of the data within the cluster. The problem of solving the minimum average residual of the data clusters is transformed into a quadratic form, and the missing data values are solved using the idea of quadratic minima. The specific algorithm is as follows.

Let B be the data set, C be the corresponding set of expression attributes, and a_{ij} be the data elements in the multidimensional data expression matrix D. Let I and J denote the subsets of B and C, respectively; then the average residuals of I and J for the specified submatrix A_{ij} are calculated as follows:

$$Z(I, J) = \frac{1}{|I \parallel J|} \sum_{i \in I, j \in J} a_{ij}\left(a_{ij} - a_{lj} - a_{iJ} + a_{IJ}\right), \quad (1)$$

where a_{ij} is the mean of row i of the submatrix, a_{iJ} is the mean of column j of the submatrix, and a_{IJ} is the mean of the submatrix. When submatrix A_{ij} satisfies $Z(I, J) \leq \delta$ (δ is the fitted value) and $0 \leq \delta$, the smaller the value of δ, the more similar the data within the corresponding submatrix is. For a bicluster matrix S containing only one missing data x, assume that the total number of rows and columns of S are m
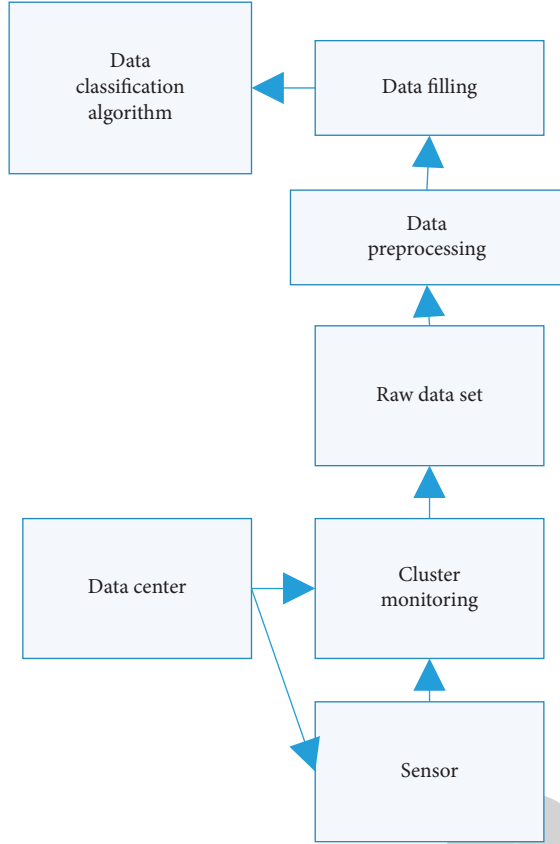
FIGURE 4: E-government big data system management architecture.

and n, respectively; the rows and columns of the missing data are p and q, respectively; the sum of all data in S except the missing data is SUM; the values of p are $(1, 2, \ldots, p-1, p+1, \ldots, n)$ [12]; and q takes values in the range $(1, 2, \ldots, q-1, q+1, \ldots, n)$, and the average residual of the matrix of data clusters is given by

$$Z(m, n) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} Z_{ij}, \qquad (2)$$

where

$$Z_{ij} = \left(a_{ij} - a_{Ij} - a_{iJ} + a_{IJ} \right)^2,$$
$$a_{ij} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} = \frac{1}{mn} (x + \text{SUM}). \qquad (3)$$

Let $\overline{A}_p$ and $\overline{B}_p$ denote the average of all data in row i and column j of S, respectively, and A_p and B_p denote the average of row p and column q containing the missing data, respectively; then its calculation formula is

$$\begin{cases} A_p = \overline{A}_p + \dfrac{x}{n}, \\ \\ B_p = \overline{B}_p + \dfrac{x}{m}. \end{cases} \qquad (4)$$

According to the previous equation, when $i = p$ and $j = q$, we get

$$Z_{ij} = \left(x - \overline{A}_p - \frac{x}{n} - \overline{B}_p - \frac{x}{m} \right) + \frac{(x + \text{SUM})^2}{mn}, \qquad (5)$$

where x is the missing data, and then its quadratic function is calculated as follows:

$$Z_{ij} = c_{ij2}x^2 + c_{ij1}x + c_{ij0}, \qquad (6)$$

where c_{ij2}, c_{ij1}, and c_{ij0} are constants.

The minimum value of $Z(m, n)$ can be calculated by combining the nature of the quadratic function of the characteristics of the minimum value. When $Z(m, n)$ is the smallest, the degree of similarity between the data is the highest, and the formula for the missing data x is

$$x = \frac{1}{(m-1)(n-1)} \sum_{i=U} \sum_{j=V} a_{ij}. \qquad (7)$$

Based on this, the filling of missing data is completed. The data that need to be processed by the E-government big data system are very large, so the proposed E-government big data system filling algorithm needs to be parallel optimized to improve efficiency. In this way, the operational efficiency of the filling algorithm for E-government big data systems can be significantly improved.

5. Experiment

In order to verify the effectiveness of the proposed algorithm for filling and classifying E-government big data systems based on a cloud computing environment, experimental analysis is conducted to verify.

5.1. Experimental Setup. In this paper, five standardized data sets, namely Abalone, CMC, Nursery, Abalone, and Yeast, were selected from the UCI machine learning database and saved as ARFF files for system testing. Basic information of the UCI data sets used is shown in Table 1.

The experiment uses distributed cloud computing on the HAD platform, with 1 high-performance PC as the master node and 19 regular PCs as slave nodes, as shown in Table 2.

In addition, the experiment will use the CloudSim simulator as the main platform for testing the filling and classification system of the E-government big data system. CloudSim will be installed, initialized, and then simulated to output the required data for testing. Data simulation flow is shown in Figure 5.

Simulation experiments can be carried out based on the above process.

5.2. Experimental Indicators. The experimental metrics for this experiment are divided into two parts: fill accuracy and classification accuracy.

5.2.1. Fill Accuracy. Since the missing data to be processed have a certain degree of dispersion and continuity, the determination of the filling accuracy requires a combination of both cases. In the case of dispersion, the true value of the filled value is judged to be the same as the value before

TABLE 1: Contents of the UCI data set.

Serial number	Data set	Number of samples	Attribute	Catalogue
1	Abalone	4,045	8	4
2	CMC	1,276	8	3
3	Nursery	11,298	8	3
4	Abalone	4,065	9	5
5	Yeast	1,439	6	4

TABLE 2: HAD platform PC configuration table.

Name	High-performance PC	Ordinary PC
CPU processor	I5-2620M	I5-2260M
Memory (GB)	8	4
Hard disk (GB)	1,024	512
Operating system	WIN7	WIN7
Operating environment	HAD-1.11	HAD-1.11
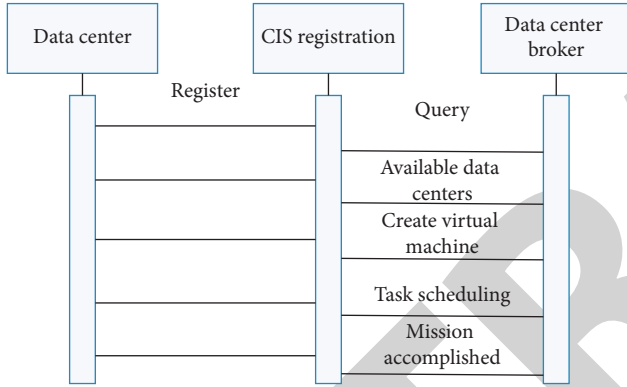
FIGURE 5: Simulation flow.

replacement, and if the true value is the same, then the filling is considered correct. The specific calculation formula is as follows:

$$P = \frac{\left(t_i + \sum_{j \in N} a g_j \left(|u_r - u| - \lambda \sqrt{S_j}\right)\right)}{n_i},$$

(8)

where t_i denotes the length of time taken to fill; a denotes the number of fills; g_j denotes the missing data finding strength factor; u_r and u denote the overall system data size before and after filling, respectively; λ denotes the variance; S_j denotes the margin of error between the filled and true values; n_i denotes the size of the data to be filled; and N denotes the set of E-government data.

5.2.2. Classification Accuracy.

The classification accuracy is a very important indicator in the evaluation of the classification effect and is calculated as follows:

$$L = \frac{\sum_{i=1} b_i}{N_a},$$

(9)

where b_i is the number of exact classifications and N_a is the total number of target classifications.
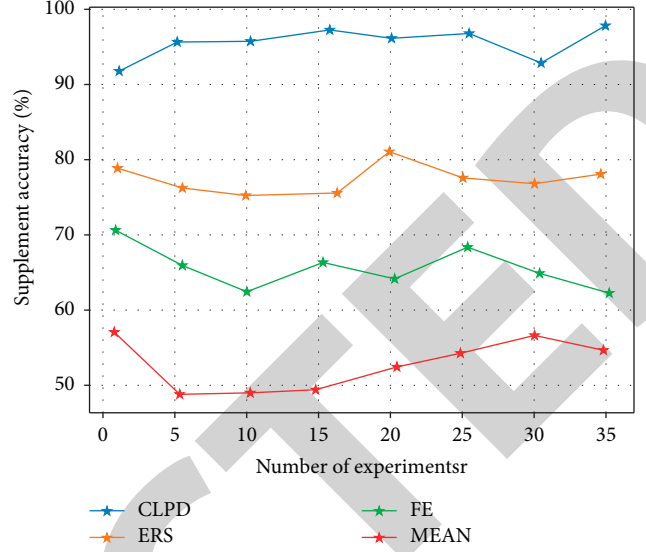
FIGURE 6: Comparison of the accuracy of filling in missing data sets.

5.3. Experimental Analysis.

In order to make the experiment more scientific and reasonable, the data set FX in the random function is used to select attribute values in the paper, and data sets with different missing rates are obtained. The traditional ways of filling missing values are the random forest algorithm MEAN method, the discrete random forest method FE method, and the weakly correlated random forest method ERS method, so this experiment combines the above three traditional methods with the cloud computing environment proposed in the paper for the E-government big data system The results are shown in Figure 6.

From the experimental results, it can be concluded that compared with the four methods, the proposed method of filling the big data system of E-government in the cloud computing environment makes full use of the information in the data set and significantly improves the accuracy rate, and the larger the amount of data, the more obvious the improvement is, compared with other methods; the proposed method in the paper has great advantages in quality and speed.

This experiment consists of two parts: the first part mainly records the changes in the number of decision forests in different datasets, establishes the change relationship between the classification model and time on the basis of this data, and establishes the change in the number of nodes in the classification model through this model. The analysis of the above data finally determines the relationship between the number of decision trees and time efficiency, as well as the impact on space efficiency and classification effect. In the second part, take the experimental results of the first part T and analyze the classification by comparing the classification of different data sets. This experiment proposes that the data sets filled by experiment one are randomly built into the classification model, and according to the decision forest algorithm, the optimal number of decision trees is 24, and 6/12/18/24/30/ decision trees are built in the established classification model. Each group was tested 10 times, and the average of the results was taken as the final result. The results of the first part of the experiment are shown in Figure 7.
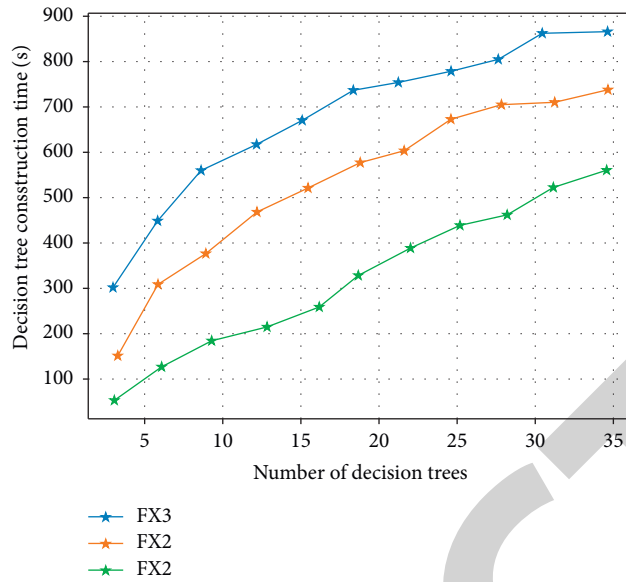
FIGURE 7: Plot of time versus number of decision trees.
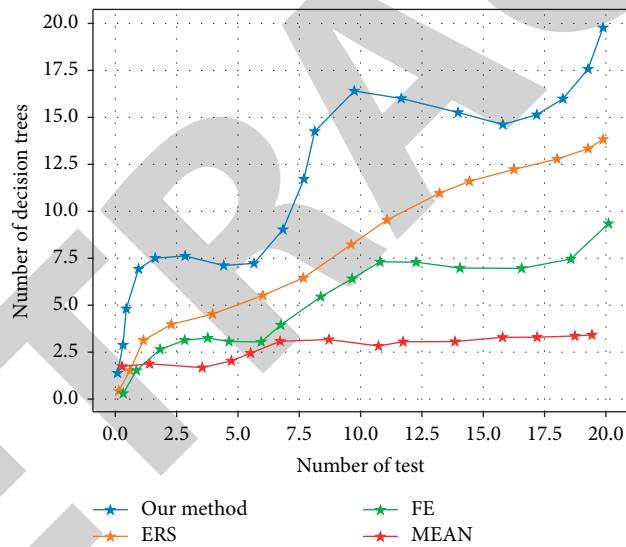
FIGURE 8: Number of nodes versus number of decision trees.

Figure 7 shows the variation of time versus the number of decision trees. From Figure 7, we can see that the number of decision trees generated by the decision forest is less than 24 in the initial period, but as the number of decision trees increases, the time gradually becomes stable, and when the number of decision trees exceeds 30, the time starts to rise.

In Figure 8, it can be concluded that the number of decision forest nodes is proportional to the number of decision trees. From Figure 8, it can be found that the classification accuracy is proportional to the number of decision trees when the number of decision trees is less than 24, and the classification accuracy is inversely proportional to the number of decision trees when the number of decision trees is more than 24. This is due to the fact that the number of decisions is too many and the decision forest algorithm

happens to have too many fits when performing the fitting; therefore, the number of decision trees is at the equilibrium point of classification efficiency when the number of decision trees is 24, and the second the results of the partial experiment are shown in Table 3.

As can be seen from Table 3, the proposed algorithm is far superior to the traditional random forest algorithm, discrete random forest algorithm, and weakly correlated random forest algorithm in terms of accuracy and reasonableness than the latter three. This is due to the fact that the method in the paper 8 takes into account the characteristics of E-government big data such as a large amount of information and wide data dimension and adopts the decision forest classification algorithm to provide a more refined decision strategy for the processing and analysis of the data

TABLE 3: Comparison of the classification accuracy of the four algorithms (%).

Number of iterations	Traditional random forest algorithm	Discrete stochastic forest algorithm	Weakly correlated random forest algorithm	Algorithm in this paper
1	78.23	76.75	80.23	95.25
2	78.64	78.47	79.45	98.34
3	86.74	80.32	81.26	94.33
4	69.04	74.85	80.54	95.25
5	71.65	79.42	83.69	96.22
6	80.59	75.73	80.83	94.34
7	75.29	76.39	79.51	96.36
8	75.84	77.54	78.84	94.86

while maintaining the updated state of the data through reconstruction processing instead of the tagged decision tree, in order to improve the accuracy of E-government big data classification.

6. Conclusions

In recent years, the development of cloud computing technology has become more and more mature and has a very important economic position in the development of the global industry. If cloud computing technology is applied to E-government big data systems, it has a very bright future, and the paper proposes a filling algorithm for E-government big data system based on cloud computing environment according to the problem of data loss that occurs when the government uses E-government big data centre system. The security and integrity of E-government data are ensured to avoid data missing. On the basis of this algorithm, a parallel optimization for filling E-government big data system in a cloud computing environment is proposed to further improve the efficiency of filling missing data; a decision forest algorithm based on discrete weak correlation is proposed, aiming at the parallel classification of E-government data for enhancing the processing capability of E-government data centres for data information; the paper further proposes an incremental update of the decision forest algorithm according to the decision forest algorithm; the optimal number of decision trees is 24.

Data Availability

The data sets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] S. Tang and M. Hanneghan, "State-of-the-Art model driven game development: a survey of technological solutions for game-based learning," *Journal of Interactive Learning Research*, vol. 22, no. 4, pp. 551–605, 2011.

[2] F. Shi, "Government data exchange platform and algorithm implementation in cloud computing environment," in *Proceedings of the 2021 4th International Conference on Information Systems and Computer Aided Education*, pp. 2401–2405, Dalian, China, September 2021.

[3] S. M. Kogan, M.-K. Lei, S. R. H. Beach et al., "Dopamine receptor gene D4 polymorphisms and early sexual onset: gender and environmental moderation in a sample of african-American youth," *Journal of Adolescent Health*, vol. 55, no. 2, pp. 235–240, 2014.

[4] K. E. Ali, S. A. Mazen, and E. E. Hassanein, "Assessment of cloud computing adoption models in e-government environment," *International Journal of Computational Intelligence Studies*, vol. 7, no. 1, p. 67, 2018.

[5] Q. Wu and Z. Li, "An adaptive service selection approach in mobile environment," in *Proceedings of the 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 122–126, Hangzhou, China, November 2012.

[6] Y. Wang, C. Fan, and Z. Wen, "A study for task time performance dynamic prediction model in cloud resource scheduling," in *Proceedings of the 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 348–353, Hangzhou, China, November 2012.

[7] J. C. P. Cheng and M. Das, "A Bim-based web service framework for green building energy simulation and code checking," *Electronic Journal of Information Technology in Construction*, vol. 19, pp. 150–168, 2014.

[8] A. Hovav, "Using scenarios to understand the frontiers of IS: fifteen years later (a postscript)," *Information Systems Frontiers*, vol. 16, no. 3, pp. 347–352, 2014.

[9] A. N. Belikov, F. Dijkstra, H. Gankema, J. B. A. N. van Hoof, and R. Koopman, "Information systems playground-the target infrastructure, scaling astro-WISE into the petabyte range," *Experimental Astronomy*, vol. 35, no. 1, pp. 367–389, 2013.

[10] V. Chang and A. Tan, "An ecosystem approach to knowledge management," *Advances in Intelligent Systems and Computing*, vol. 172, pp. 25–35, 2013.

[11] C. H. Cao, Y. N. Tang, D. Y. Huang, W. M. Gan, and C. Zhang, "IIBE: an improved identity-based encryption algorithm for WSN security," *Security and Communication Networks*, vol. 2021, Article ID 8527068, 8 pages, 2021.

[12] J. Yan, X. Yun, and Z. Wu, "A novel weighted combination technique for traffic classification," in *Proceedings of the 2012 IEEE 2nd International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 757–761, Hangzhou, China, November 2012.

[13] D. Wang, A. Fan, and L. Xiong, "Design and implementation of mirror traffic network simulation framework based on Delta3D," in *Proceedings of the IEEE 2nd International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 1157–1160, Hangzhou, China, November 2012.

Research Article

Application of Machine Learning in the Reliability Evaluation of Pipelines for the External Anticorrosion Coating

Zhifeng Zhao [iD],[1] **Mingyuan Chen,**[2] **Heng Fan** [iD],[1] **and Nailu Zhang**[1]

[1]*School of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China*
[2]*Dept of Mechanical, Industrial and Aerospace Engineering, Concordia University, Montreal H3G2W1, Canada*

Correspondence should be addressed to Zhifeng Zhao; zfzhao@xsyu.edu.cn

The purpose of this research is to enhance the analysis of the reliability status for external anticorrosive coatings. With the limitation and insufficiency of the static evaluation method, we study and construct an evaluation method of dynamic reliability for the anticorrosive layer, integrating the trend analysis of the Markov chain and the set pair theory. This method is implemented by the machine learning software of PyCharm community edition, based on *Python* language. The algorithm utilizes the connection degree in the set pair theory to determine the risk levels of the anticorrosive coating systems. According to the characteristics of the dynamic change of the anticorrosive layer with time, we built the mathematical evaluation model by combining it with the nonaftereffect property of the Markov chain. Therefore, we can make a dynamic and useful analysis for the reliability grade of the anticorrosive coating and assess the effectiveness grade of the changed reliability for the anticorrosive coating after some time. This method can effectively evaluate the reliability level of the anticorrosion coating through the example of big data of detection points. Under national standards, we provide the theoretical basis for pipeline maintenance within detection cycle requirements.

1. Introduction

The reliability evaluation of the pipeline external anticorrosive coating is a significant component of pipeline integrity management, which is related to the safety of pipeline energy transportation [1]. The current safety evaluation method mostly belongs to static evaluation. Because of the dynamic characteristics of the external anticorrosive coating, the static evaluation method cannot meet the dynamic requirements of the system in some aspects. Their numerical conclusions are often too simple to express danger. The poor expression of risk tendentiousness and dynamics is not conducive to providing decision-making services to managers [2, 3].

According to the research of reliability analysis methods at home and abroad, there are two main categories of external anticorrosion coating for pipelines: the evaluation method of direct safety and the evaluation method of mathematical theory analysis [4]. The evaluation methods of

direct safety include the method of fault tree assessment, the safety checklist, the risk matrix diagram, the event tree analysis, the direct assessment method of external corrosion, and the manual of pipeline risk management. They belong to qualitative or semiquantitative methods. Generally, they can only give the risk levels or relative values, such as high, medium, or low. They have no unit and dimensional values or accident probability and have certain limitations. The analysis evaluation methods of applying mathematical theory include the analytic hierarchy process, the model method of BP neural network, the grey theory, and the fuzzy evaluation method. The model method of the BP neural network is to use the autonomous learning of neural network models and nonlinear approximations to obtain knowledge. It puts forward the ideas of serial-parallel processing and self-fault tolerance. However, the BP algorithm is only an improved method of nonglobal partial search. It requires that the solution is indeed the global extremum of a nonlinear function, which may cause the algorithm to fall

into the local extremum and lead to training failure [5]. The analytic hierarchy process treats complex problems as systems. It decomposes the multiindex or constraint into various levels so as to evaluate and analyze the problems at various levels and find out the decision-making method of multiindex scheme optimization. However, if the amount of index data is too large, it is difficult to solve the eigenvalues and eigenvectors of the matrix, and the weight cannot be determined, resulting in the failure of the analysis. In the process of scheme formulation, the better one can only be selected from the known hierarchical schemes, but the shortcomings of the known schemes cannot be found, and the improvement methods cannot be put forward. So it cannot provide a new scheme for decision-making for evaluation. The fuzzy evaluation method of grey theory is to analyze data by using fuzzy mathematics and the grey system analysis method. However, the algorithms for parameter threshold and correlation degree in its membership function have obvious defects, which will have a great impact on the accuracy of the results [6, 7].

Aiming at the limitations and shortcomings of static evaluation, the dynamic evaluation of the deterioration trend of the anticorrosive coating based on the set pair (S.PA) and Markov chain theory is proposed [8, 9]. By using the connection degree of S.PA to determine the status of each deterioration grade in the anticorrosive coating, we established the reliability evaluation model of the degradation trend level of the anticorrosion layer [10, 11]. Combining the nonaftereffect and homogeneous of the Markov chain, the safety effectiveness of the anticorrosive coating varying at intervals of time is obtained according to the transfer matrix and modification [12, 13]. Through the above algorithm, we use the machine learning software of the PyCharm community edition based on the *Python* language to predict and intelligently analyze the detection data.

2. Methods

2.1. Set Pair Theory. Set pair theory uses connection degree to represent the connection component of two sets under the background of a certain problem. The connection degree is expressed by the normalized connection number. That is, $U = a + bi + cj$. "a" denotes the identity degree. "b" denotes the difference degree. That is, the difference is uncertain, and the characteristics are neither identical nor opposite in the two sets. "i" is a marker symbol or a difference coefficient. "c" is called the opposition degree. "j" is a marker symbol or a degree of opposition coefficient [14]. By using the connection formula of S.PA, we can use this method to describe the deterioration status of pipeline anticorrosive coating [15].

2.2. Markov Chain. It is a method proposed and named by the Russian mathematician Markov to analyze the stochastic process of things. It is different from other methods such as variance analysis and interval estimation. It has the characteristics of no aftereffect. Given the corrosion status of the pipeline anti-corrosive coating, the future corrosion status is only related to the corrosion status at the time of detection, which has the characteristics of the Markov chain. Therefore, the corrosion degradation process can be viewed as a Markov chain [16, 17].

For the pipeline coating, the transition probability between the degradation states of the coating is only related to the corrosion state. It has nothing to do with the time n of detection. Therefore, the corrosion state process can be regarded as a homogeneous Markov chain [18]. That is, in the case of any state $i, j \in S = \{1, 2, \ldots, n\}$ (state space), the transition probability $P_{ij}^{(n)}$ of the Markov chain is only related to state i and j. $P_{ij}^{(1)}$ is used to represent one-step transition probability, that is,

$$P_{ij}^{(1)} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix}, \tag{1}$$

where $P^{(n)} = (P_{ij}^{(n)})$ is n-step transition probability matrix, and $P_{ij}^{(n)} \geq 0$, $\sum_{k \in S} P_{ij}^{(n)} = 1$. The n-step transition probability $P_{ij}^{(n)}$ has the following properties [19]:

$$\begin{aligned} P_{ij}^{(n)} &= \sum_{k \in S} P_{ik}^{(l)} P_{kj}^{(n-l)}, \\ P_{ij}^{(n)} &= \sum_{k_1 \in S} \cdots \sum_{k_{n-1} \in S} P_{ik_1} P_{k_1 k_2} \cdots P_{k_{n-1} j}, \\ p^{(n)} &= p p^{(n-1)}, \\ p^{(n)} &= p^n. \end{aligned} \tag{2}$$

At present, there are four methods for calculating the transition probability matrix, such as the solving method of the inverse matrix, the regression analysis method, the empirical judgment method, and the statistical analysis method.

2.2.1. Statistical Analysis Method. The statistical analysis method calculates the percentage of one to another state in a certain period under the condition of data accumulation for many years. That is to say, the number m_i in the hypothesis "i" state is calculated. From "i" to "j" state, the transition number m_j is also calculated. The transition probability $P_{ij}^{(1)}$, $P_{ij}^{(1)} = m_j/m_i$ is used as the element of the matrix of transition probability, and then the matrix of transition probability is constructed in turn.

2.2.2. Empirical Judgment Method. The empirical judgment method requires using the engineering experience to determine the transfer probability matrix. It is subjective and one-sided. The magnitude of each probability value is estimated by the subjective human. This method needs several years of verification and gradual revision before it tends to the actual state of the transfer law.

2.2.3. Inverse Matrix Method. The inverse matrix method considers the inverse matrix to solve the transition probability matrix. It requires a few detection data. With the addition of new data, the transfer probability matrix can be amended to make the judged data results more reasonable and accurate. The method steps are as follows:

Assuming that the system has three states, the initial vector is $U_1 = \{a_1, b_1, c_1\}$, and the matrix of transition probability is

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ p_{21} & p_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}. \tag{3}$$

Then the next year's value is $U_2 = \{a_2, b_2, c_2\}$ and can be expressed as

$$U_1 \times P = U_2. \tag{4}$$

The equation is

$$\begin{aligned} a_1 P_{11} + b_1 P_{21} + c_1 P_{31} &= a_2, \\ a_1 P_{12} + b_1 P_{22} + c_1 P_{32} &= b_2, \\ a_1 P_{13} + b_1 P_{23} + c_1 P_{33} &= c_2, \\ P_{11} + P_{12} + P_{13} &= 1, \\ P_{21} + P_{22} + P_{23} &= 1, \\ P_{31} + P_{32} + P_{33} &= 1. \end{aligned} \tag{5}$$

In the same way, the values of $U_3 = \{a_3, b_3, c_3\}$ and $U_4 = \{a_4, b_4, c_4\}$ in another two years can be listed into eight equations of the same form. The total set has twelve equations. The transition probability parameter solution has nine parameters. Therefore, four years of data can be used to solve the elements of the transition probability matrix in the system of equations, that is,

$$C \times P = D, \tag{6}$$

$$C = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \tag{7}$$

$$D = \begin{pmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{pmatrix}. \tag{8}$$

The transition probability matrix is

$$P = C^{-1} \times D. \tag{9}$$

2.2.4. Regression Analysis. Regression analysis is based on the assumption that the transfer probability distribution obeys a certain distribution law. For example, it follows a normal distribution and then obtains the probability distribution of the estimated value. The first step requires establishing the regression equation between the variables (regression accuracy W). The next step is to determine M_i (the median value of state i) into the regression equation and inversely calculated t_i (number of years). Then t_{i+1} is substituted into the regression equation, calculated the expected value for the next year and assumed that the transition probability also obeys the normal distribution. According to the expected value, we determine the standard deviation (accuracy of the regression equation) and the discrete state distribution. So, we determine the probability distribution of each state, and construct the transition probability matrix P. From the above calculation steps, we observed that the results are to be verified by assuming that the distribution law of transition probability obeys normal distribution. Then the regression model is checked, and the errors are calculated and predicted.

Establishment steps of a dynamic assessment model for the deterioration tendency grade of external anticorrosion coating.

According to the evaluation standard of external anticorrosion coating, we select reasonable performance parameters that can reflect the damage and deterioration level of the anticorrosion coating and define the deterioration level of the external anticorrosion coating. We establish the evaluation model of connection degree by using the set pair theory.

The initial state distribution of the anticorrosion coating system is established by using the detection data of the external anticorrosion coating in different years, and the solution method of the appropriate transfer probability matrix is selected to calculate the transfer probability matrix.

Combined with the properties of the Markov chain, we establish the dynamic evaluation model of the Markov chain of the set pair theory by using the transition probability matrix and initial state distribution. Therefore, we evaluate the distribution of deterioration states of the external anticorrosion coating [20].

Application of dynamic assessment model for the deterioration tendency grade of external anticorrosive coating.

The actual detection methods and data of gas and oil pipelines in the project site allow us to select the breakage points density classification of the anticorrosive coating as the evaluation index. At present, the methods and technologies to judge the damage point of anticorrosion coating are as follows: numerical direct reading method, radiation distance method, statistical graphics method, voltage-controlled vibration frequency method, cursor observation method, and magnetic field drop method [21, 22]. With the actual requirements of maintenance measures, the deterioration level of external anticorrosive coating is divided into three levels: the first level is classified as the excellent and good status of the damage point density; the second level is a normal status of the damage point density; and the third level is classified as a bad status of the damage point density [23, 24]. Table 1 shows the classification standard for the degradation of the external anticorrosive coating. "Bp" is the density value of the breakage point of the anticorrosive coating. The minimum distance between two adjacent damaged points is not more than twice the central burial depth of the pipeline, which can be marked as one.

TABLE 1: Classification standard for the degradation of the external anticorrosive coating.

Type of external anticorrosive coatings	Degradation status classification			
	Level I		Level II	Level III
Asphalt anticorrosive coating, anticorrosive insulation layer of rigid polyurethane foam	1 (excellent)	2 (good)	3 (normal)	4 (Bad)
	Bp ≤ 0.2	0.2 < Bp < 1	1≤Bp ≤ 2	Bp > 2

According to the connection degree of the S.PA, a definition is introduced: $U_m = a + bi + cj$, m = (1, 2, ..., n) (natural number). "a" represents the first-class grade of excellent and good anticorrosive coating in the whole pipeline. "b" indicates the second-class grade of normal anticorrosive coating in the whole pipeline. "c" shows the third-class grade of bad anticorrosive coating in the whole pipeline. "i" represents the marker of the normal grade of pipeline. "j" indicates the marker of the bad grade and opposition to the excellent and good grade. As an example of the Bp data of breakage points of asphalt anticorrosive coating for the 6909-meter pipeline, we calculate that the pipeline length with excellent and good anticorrosive coating is 4421m, so $a = 4421/6909 = 0.6399$. Similarly, the pipeline length in the second grade is 2211m, so $b = 2211/6909 = 0.3200$. The pipeline length in the third-grade is 277m, so $c = 277/6909 = 0.0401$. Table 2 describes the classification status summary of inspection data collated in different periods for the anticorrosive coating (detection cycle every three years). Therefore, the initial state distribution model for the anticorrosive coating pipeline is $U_1 = 0.6399 + 0.3200i + 0.0401j$. In other years, the state distribution models of the anticorrosive coating pipeline are as follows:

$$
\begin{aligned}
U_2 &= 0.4198 + 0.4101i + 0.1701j, \\
U_3 &= 0.3198 + 0.4100i + 0.2702j, \\
U_4 &= 0.2698 + 0.4000i + 0.3302j, \\
U_5 &= 0.2451 + 0.3901i + 0.3648j.
\end{aligned}
\tag{10}
$$

We observed the state model that the deterioration of coating becomes more and more obvious. That is, the changing trend of the good grade is from 0.6399 to 0.2451. With the increase of service years, the good grade becomes smaller. The changing trend of the normal grade is from 0.3200 to 0.3901. By comparing the indexes nearly four times, we found that the changing trend of normal grade is about 0.40, which conforms to the basic change law of the pipeline life cycle. The changing trend of the bad grade is from 0.0401 to 0.3648, and the 4th deterioration state of coating has begun to be greater than that of the good grade (0.3302 > 0.2698). The trend of aging is obvious.

The same conclusion can be drawn from the set-pair tendency. When the set-pair tendency is used for the reliability system evaluation, the set-pair tendency represents the reliability tendency. Its reliability analysis is based on the hierarchical relationships in the connection number. Taking the ternary connection number as an example, it indicates the degree to which the relationship between (a, b) and (c)

affects the deterioration development. That is, $\partial c = c/(a + b)$, which indicates the dynamic development of the deterioration direction. In the example, the deterioration level of external anticorrosion coating for the pipeline has developed from $0.0401/(0.6399 + 0.3200) = 0.0418$ to $0.3648/(0.2451 + 0.3901) = 0.5743$, and the degree of deterioration is obviously large.

According to the analysis of the data characteristics of Table 2, we selected the inverse matrix method to solve the transition probability matrix. Therefore, the machine learning software of PyCharm community edition is used for data analysis. We construct the matrix C_1 and D_1 by the four distribution vectors corresponding to the detection data of the previous four times. That is,

$$
C_1 = \begin{bmatrix} 0.6399 & 0.3200 & 0.0401 \\ 0.4198 & 0.4101 & 0.1701 \\ 0.3198 & 0.4100 & 0.2702 \end{bmatrix},
$$
$$
D_1 = \begin{bmatrix} 0.4198 & 0.4101 & 0.1701 \\ 0.3198 & 0.4100 & 0.2702 \\ 0.2698 & 0.4000 & 0.3302 \end{bmatrix}.
\tag{11}
$$

According to equation (2)–(9), we calculate the transfer probability matrix P_1 through python programming based on the previous four detection data,

$$
P_1 = \begin{bmatrix} 0.5642 & 0.3684 & 0.0674 \\ 0.1756 & 0.5112 & 0.3132 \\ 0.0643 & 0.2686 & 0.6671 \end{bmatrix},
\tag{12}
$$

with the properties of the Markov chain, we evaluate the deterioration value of anticorrosive coating in June 2021,

$$
U_5' = \begin{bmatrix} 0.2437 & 0.3926 & 0.3637 \end{bmatrix}.
\tag{13}
$$

Compared with the fifth detection value in August 2021, there exists a certain deviation. Therefore, we modify the transition probability matrix.

Firstly, we establish four distribution vectors corresponding to the last four detection data. We construct the matrices C_2 and D_2, which is

$$
C_2 = \begin{bmatrix} 0.4198 & 0.4101 & 0.1701 \\ 0.3198 & 0.4100 & 0.2702 \\ 0.2698 & 0.4000 & 0.33\ 02 \end{bmatrix},
$$
$$
D_2 = \begin{bmatrix} 0.3198 & 0.4100 & 0.2702 \\ 0.2698 & 0.4000 & 0.3302 \\ 0.2451 & 0.3901 & 0.3648 \end{bmatrix}.
\tag{14}
$$

TABLE 2: Summary table of inspection data for the coating classification status.

Number of detections	Level		
	Level I (excellent and good)	Level II (normal)	Level III (bad)
1 (2009.06)	0.6399	0.3200	0.0401
2 (2012.07)	0.4198	0.4101	0.1701
3 (2015.07)	0.3198	0.4100	0.2702
4 (2018.06)	0.2698	0.4000	0.3302
5 (2021.08)	0.2451	0.3901	0.3648

Secondly, according to equation (2)–(9), we calculate the transfer probability matrix P_2 through python programming based on the last four detection data,

$$P_2 = \begin{bmatrix} 0.6223 & 0.2658 & 0.1119 \\ 0.0921 & 0.6587 & 0.2492 \\ 0.1223 & 0.1663 & 0.7114 \end{bmatrix}. \qquad (15)$$

Finally, P_1 is used to take the mean value of P_2 to modify the transition probability matrix, and P_{mod} is calculated by the python program,

$$P_{mod} = \begin{bmatrix} 0.5933 & 0.3171 & 0.0896 \\ 0.1339 & 0.5849 & 0.2812 \\ 0.0933 & 0.2175 & 0.6892 \end{bmatrix}. \qquad (16)$$

According to the modification of the transfer probability matrix, the deterioration value of the anticorrosive coating by the *Python* program algorithm in August 2024 is predicted as follows:

$$U_6 = 0.2317 + 0.3852i + 0.3831j. \qquad (17)$$

Similarly, the deterioration value of the anticorrosive coating in August 2027 is predicted as follows:

$$U_7 = 0.2248 + 0.3821i + 0.3931j. \qquad (18)$$

The above flow chart of the *Python* program is shown in the *Python* flow chart of the set pair-Markov algorithm in Figure 1. Figure 2 is the python calculation value of the set pair-Markov algorithm for the pipeline's external anticorrosion coating. The data vector iterates through the set pair-Markov algorithm until the product of the vector remains unchanged, and outputs the result. Figure 3 is the iterative value diagram for predicting the deterioration of the external anticorrosion coating of the pipeline against the set pair-Markov matrix vector. The red dot represents the iterative value distribution of pipeline tendency in excellent and good condition. The green dot represents the iterative value distribution of pipeline tendency in the general state. The blue dot represents the iterative value distribution of pipeline tendency in a poor state.

3. Results

In August 2024, the first class of the anticorrosion layers will reach 23.17%, the second class will reach 38.52%, and the third class will reach 38.31%. In August 2027, the

deterioration resulting in anticorrosion coating will have 22.48% in the first class, the second class will have 38.21%, and the third class will have 39.31%. From 2024 to 2027, the deterioration tendency further increases, and the poor state of the anticorrosion coating is nearly twice that of the good state. The degree of deterioration is larger. Through program calculation, it is found that the transition probability matrix P will basically not change and the limit matrix B will appear when $m \geq 11$. That is, $B = \begin{bmatrix} 0.2173 & 0.3780 & 0.4047 \\ 0.2173 & 0.3780 & 0.4047 \\ 0.2173 & 0.3780 & 0.4047 \end{bmatrix}$. For steady-state, it is worth noting that the product of the transition matrix and steady-state must be steady-state again. When the limit matrix B is obtained, the steady-state distribution probability of the quality status of the anticorrosion coating for the whole pipeline can be obtained. The steady-state distribution probability $\pi = [0.2173, 0.3780, 0.4047]$. It shows that the steady-state distribution probability is independent of the initial vector U4 and only related to the limit matrix B. It can be seen that the quality distribution will tend to be stable when the inspection times of the anticorrosion coating for the pipeline are enough. The proportion of pipeline sections with a quality grade of excellent and good is 21.73%. The proportion of pipeline sections with a general quality grade is 37.80%. The proportion of pipeline sections with a poor-quality grade is 40.47%. The excellent and good condition is maintained at one fifth of the pipeline. The other two conditions account for two fifths each. The deterioration tendency is obvious. We need to take maintenance measures to prevent the deterioration of the anticorrosion coating.

4. Discussion

With the development of the corresponding detection technology for the pipeline, the accuracy and statistical management of coating detection data will be more meticulous in the improvement and standardization of the relevant standards. Meanwhile, the methods of modifying the transfer matrix are also increasing, such as the arithmetic average method, weighted arithmetic average method, geometric average method, weighted geometric average method, square average method, and the harmonic average method. The transfer probability matrix P can be updated and revised in time, making the evaluation results more in line with the actual situation on the spot. So, we will grasp more accurately the deterioration trend of the anticorrosion layer. It has great significance for prolonging the service life of pipelines.

```
                    Input 3x3 matrix C1 and D1 vectors according to the first four groups of detection data of external
                                                    anti-corrosion coating
```

$$P_1 = C_1^{-1} \times D_1$$

| Calculate the transition probability matrix P1 | ⟸ | Calculate the new transition probability matrix P2 |

Take the mean value for correction

Input the fifth detection data matrix vector 'U4' or 'U5'value ⟹ Set pair-Markov algorithm

$$P_2 = C_2^{-1} \times D_2$$

Compare the actual detection data in the prediction year of matrix vector of the prediction, and decide whether to modify the transition probability matrix

Y ⟹ Input 3x3 matrix C2 and D2 vectors according to the new four groups of detection data of external anti-corrosion coating

N

Iterative calculation, and output the prediction results and vector value diagram in the corresponding service time

FIGURE 1: The python flow chart of the set pair-Markov algorithm.
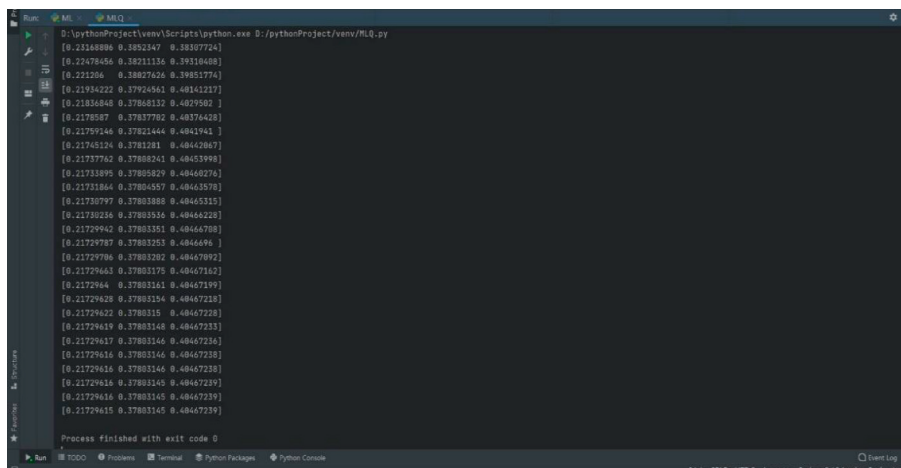
FIGURE 2: The *Python* calculation value of the set pair-Markov algorithm for the external anticorrosion coating of the pipeline.
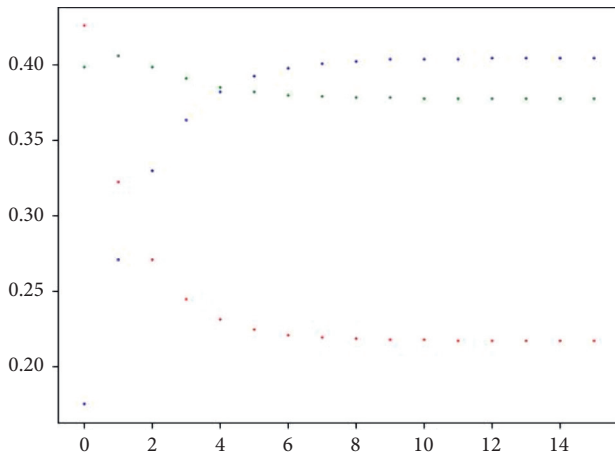
Figure 3: The iterative value for predicting the deterioration of the external anticorrosion coating of pipeline against the set pair-Markov matrix vector.

5. Conclusions

Through the case analysis, we evaluate the deterioration status of the anticorrosive coating in 2024 and 2027 and judge the safety grade of the anticorrosive coating comprehensively. The evaluation results can reflect the reliability status of the anticorrosive coating of pipelines in a timely, objective, and comprehensive manner. It makes up for the limitations and shortcomings of static evaluation. It also provides a decisive theoretical basis for pipeline maintenance.

If the corresponding maintenance measures are taken for the normal and bad anticorrosion coatings, the reliability and quality level of the anticorrosion coating will be improved. But the deterioration of the anticorrosion coating will continue to deteriorate without any maintenance measures. This is a dynamic characteristic of the damage distribution of the external anticorrosion coating. Combining with the machine learning software of PyCharm community edition based on the *Python* language, the intelligent detection data analysis is carried out by taking advantage of its simplicity, wide application fields, high development efficiency, and comprehensive class library. Through the set pair theory with the Markov chain, we can effectively analyze the change, and evaluate the distribution of deterioration trend of pipeline coating. According to the actual situation of pipelines, each pipeline company can guide and formulate the corresponding evaluation and detection times for the three to eight-year cycles in the national standards and integrate them. The optimization of pipeline anticorrosive coating is guaranteed under reasonable safety and economic conditions. It also provides a basis for the pipeline maintenance of formulating scientific and efficient schemes.

Data Availability

The raw data is obtained through the pipeline breakage points collected every three years in the third-party detection department of PetroChina Changqing Oilfield Company. The raw data required for these findings cannot be shared at this time as the data also forms part of an ongoing study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

Zhifeng Zhao (first author) contributed to the methodology, software, investigation, data analysis, and writing of the original draft. Mingyuan Chen contributed to writing the review and investigation. Heng Fan contributed to the data curation and resources. Nailu Zhang contributed to the investigation and resources.

Acknowledgments

References

[1] Standards Press of China, Standardization for External Corrosion Control of Steel Pipelines, GB/T 21447-2008, Standards Press of China, Beijing, China, 2008.

[2] Z. F. Zhao, H. Wen, and W. X. Gao, "Data mining and knowledge decision in the integrity management of long-distance pipeline," *Journal of Xi'an Shiyou University (Natural Science Edition)*, vol. 31, no. 4, pp. 109–114, 2016.

[3] Y. Y. Cui, M. M. Zhang, and Y. Z. Liu, "The evaluation criterion for external coating performance of buried pipelines," *Applied Chemical Industry*, vol. 44, no. 12, pp. 2312–2316, 2015.

[4] Z. F. Zhao, M. Y. Chen, and H. Fan, "Safety evaluation of comprehensive dynamic for the external anti-corrosive system of long-distance pipeline," in *Proceedings of the International Petroleum and Petrochemical Technology Conference 2020*, pp. 321–326, Beijing, China, 2021.

[5] X. M. Han, H. Miao, and Z. Wang, "Pipeline integrity prediction method based on big data and neutral network," *Oil & Gas Storage and Transportation*, vol. 34, no. 10, pp. 1042–1046, 2015.

[6] Z. S. Luo, A. R. Bi, and X. W. Wang, "Research into corrosion of wet natural gas gathering and transportation pipeline based on gray support vector machine," *China Safety Science Journal*, vol. 26, no. 2, pp. 85–90, 2016.

[7] A. Vaezi and M. Verma, "Railroad transportation of crude oil in Canada: developing long-term forecasts, and evaluating the impact of proposed pipeline projects," *Journal of Transport Geography*, vol. 69, pp. 98–111, 2018.

[8] Petroleum Industry Press, Oil and Gas Pipeline Risk Assessment Method Part I Semi-Quantitative Assessment Method, SY 6891.1-2012T, Petroleum Industry Press, Beijing, China, 2012.

[9] Z. F. Zhao, H. Wen, and W. X. Gao, "Integrated evaluation of the soil corrosion in pipeline in contrary identical discrepancy model," *Journal of Xi'an University of Science and Technology*, vol. 37, no. 3, pp. 352–357, 2017.

[10] Z. F. Zhao, H. Wen, and J. Guo, "Set pair theory and clustering analysis method for coal and gas outburst prediction and its application," in *Proceedings of the ICEEP 2015*, pp. 784–790, Shenzhen, Guangdong, China, 2015.

[11] J. Tejedor, J. Macias-Guarasa, H. F. Martins, S. Martin-Lopez, and M. Gonzalez-Herraez, "A Gaussian Mixture Model-Hidden Markov Model (GMM-HMM)-based fiber optic surveillance system for pipeline integrity threat detection," in *Proceedings of the 6th International Conference on Optical Fiber Sensors*, Lausanne, Switzerland, September 2018, https://opg.optica.org/abstract.cfm?URI=OFS-2018-WF36Optical Fiber Sensors 2018.

[12] A. Valor, F. Caleyo, L. Alfonso, J. C. Velázquez, and J. M. Hallen, "Markov chain models for the stochastic modeling of pitting corrosion," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–13, 2013, to be published.

[13] S. Yusof, N. M. Noor, N. Yahaya, and A. S. A. Rashid, "Markov chain model for predicting pitting corrosion damage in offshore pipeline," *Asian Journal of Scientific Research*, vol. 7, no. 2, pp. 208–216, 2014.

[14] K. Q. Zhao, *Set Pair Analysis and its Preliminary Application*, pp. 9–24, Zhejiang Science and Technology Press, Hangzhou, Jiangsu, China, 2000.

[15] Z. F. Zhao, H. Wen, and W. X. Gao, "A pipeline soil corrosion comprehensive prediction method using multi-method integration," *International Journal of Simulation. Systems, Science and Technology*, vol. 17, no. 26, pp. 19.1–19.8, 2016.

[16] A. L. A. Mubarak, A. Bouferguene, and Y. Mohamed, "Random generation of industrial pipelines' data using Markov chain model," *Advanced Engineering Informatics*, vol. 38, no. 10, pp. 725–745, 2018.

[17] Z. F. Zhao, D. Wu, G. W. Gao, and H. Fan, "Study on evaluation model of rate grade of soil corrosion for pipeline," in *Proceedings of the 2021 3rd International Conference on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP)*, pp. 399–402, Xi'an, China, July 2021.

[18] C. I. Ossai, B. Boswell, and I. Davies, "Markov chain modelling for time evolution of internal pitting corrosion distribution of oil and gas pipelines," *Engineering Failure Analysis*, vol. 60, no. 2, pp. 209–228, 2016.

[19] M. Li and X. L. Wang, "Detecting technology and assessment of underground oil-gas pipeline coating deficiency," *Contemporary Chemical Industry*, vol. 42, no. 7, pp. 980–983, 2013.

[20] R. J. Wang and T. Y. Wang, "Application of grey Markov chain model in corrosion forecast of buried oil and gas pipelines," *Journal of Safety Science and Technology*, vol. 11, no. 4, pp. 102–106, 2015.

[21] Standards Press of China, Inspection of Corrosion Protection Engineering for Buried Steel Pipelines, GB/T 19285-2014, Standards Press of China, Beijing, China, 2014.

[22] B. L. Sha, "Collection and transmission of corrosion prevention parameters and performance assessment of corrosion prevention in buried pipeline," M.S. thesis, Dept. Petroleum. Eng., NorthEast Petroleum University., Daqing, China, 2013.

[23] Petroleum Industry Press, Pipeline Integrity Management Specification Part 1: General Provisions, Q/SY 1180.1-2009. Petroleum Industry Press, Beijing, China, 2009.

[24] S. J. Peng, H. P. Tang, and Y. P. Ding, "Study on characteristics of safety assessment model on pipeline corrosion," *Journal of Safety Science and Technology*, vol. 3, no. 3, pp. 172–178, 2015.

Research Article

Countermeasures for Urban Traffic Congestion in China from the Perspective of System Dynamics

Wenjing Zhang (ID)

School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Wenjing Zhang; 14113103@bjtu.edu.cn

In recent years, urban traffic congestion has seriously affected the healthy development of urbanization in China. And many measures to combat congestion have had little effect . The purpose of this paper is to find out the most reasonable and sustainable measures to control traffic congestion . Based on the theory of system dynamics, this study constructs a model of the formation mechanism of urban traffic congestion in China, and analyzes the thinking error of the traditional strategy of "building roads to eliminate traffic congestion" This study includes the current policy measures to control traffic congestion in the system dynamics model and analyzes the influence of each measure on the formation mechanism of urban traffic congestion. Then, it critiques the unsustainability of rigid policies, such as the vehicle number limit and the "similar road building to control traffic congestion" policies. This study reveals that of the five policies adopted by the government to alleviate traffic congestion, and come to the conclusion: the "sparse block collocation" policy is the most sustainable and fundamental congestion control measure. To achieve efficient traffic congestion control and support the healthy development of urbanization in China in the future, the government should increase the balance of infrastructure investment to improve the slow environment of public transport, adhere to public transport-oriented land development policies, raise the cost of motor vehicle travel, and promote urban traffic.

1. Introduction

In recent years, the problem of urban traffic congestion in China has become increasingly serious. Traffic congestion first appeared in megacities such as Beijing, Shanghai, and Guangzhou, and quickly spread to smaller cities. Urban traffic congestion not only delays people's normal travel and damages their physical and mental health [1], but also reduces the efficiency of urban economic life and energy use and intensifies greenhouse gas emissions, which are serious threats to the healthy development of urban areas in China. Therefore, it is urgent to put forward and implement reasonable and sustainable measures to control the congestion. And the purpose of this study is to identify the most reasonable and sustainable treatment measure. Discussions about urban traffic congestion in academia and society have included rail transit, intelligent transportation, citizen quality, development stage, population, bus, and system theories. These theories have been used to determine the causes of urban traffic congestion from their specific fields

and personal experiences to provide corresponding countermeasures. For example, Minghua and Zhenghe studied techniques to solve traffic problems in big cities and found that effective countermeasures should be taken to adjust the urban traffic volume from the two principles of reasonably reducing traffic occurrence and improving traffic efficiency as much as possible [2]. Yuanqing et al. studied institutional causes and solutions of traffic problems in China's big cities and examined the role of urban traffic's sustainable development from the perspective of the law of urban traffic development [3]. Xianglin et al. explored coupled mapping following model and traffic congestion control based on an intelligent transportation system [4]. As Reported by the Nanfang Daily, National People's Congress deputies stated that a large number of cars was not the root cause of congestion [5]. Feng et al. studied how to regulate the population scale in megacities, and believed that populations in the main urban areas should be evacuated, infrastructure investment should be increased, and urban industrial structures and economic development speed should be

adjusted [6]. Taking Nanjing as an example, Xianteng et al. explored the contradiction between the supply and traffic demand in big cities and development countermeasures and believed that the fundamental root of traffic congestion was the imbalance between the traffic supply and demand [7]. The above studies suggested the causes of urban traffic congestion from different perspectives and proposed the corresponding governance measures.And the findings of the previous research provides reference and inspiration for this study. However, these studies lack critical synthesis and have not provided any detailed analysis of urban traffic congestion from a systematic viewpoint. Therefore, from the perspective of system dynamics (SD), the current study provides a detailed analysis of urban traffic congestion and its governance measures, which is innovative to a certain extent and is of certain significance to this field.

Under the pressure of traffic congestion, the central government and local governments in China have continuously discussed and introduced a series of anti-congestion policies. The Ministry of Transport has stated that efforts should be made to build an urban public transport system and increase investment in bus infrastructure and hub stations to solve the problem of urban congestion [8]. Beijing has also issued the Opinions on Further Promoting the Work of Easing Traffic Congestion in the Capital, which formulates comprehensive measures from the following five aspects: improving urban planning, speeding up road traffic infrastructure construction, prioritizing public transportation development, improving the non-motor vehicle traffic system, and strengthening the management and scientific management of motor vehicles [9]. In response to congestion and with the aim of achieving a speed of no less than 25 km/h on central urban roads within 5 years, Guangzhou has issued 30 ordinances, including raising parking fees, studying the levy of traffic congestion fees, and studying the prohibition of foreign cars into the city [10].

Traffic congestion, as a trending urban problem, has also been applied in SD studies in recent years. For example, Wang et al. studied the SD model of an urban traffic system and its application. Building on the SD method, the authors established a model of the urban traffic system based on system structure analysis and causality feedback analysis [11]. Furthermore, Shuang conducted an empirical study on the evolution mechanism of urban traffic structures based on SD [12]. Finally, Yimei and Yi studied previous SD model simulations of urban traffic congestion, analyzed the formation mechanism of urban traffic congestion, and established an SD model of urban traffic congestion [13].

Unfortunately, the existing SD models of traffic congestion in the aforementioned studies are characterized by macroscopic complexity, numerous internal module elements, and attempts to cover all aspects of the urban economy and society. The main defects are that the setting of the traffic demand generation module is too simple, and some important factors (such as hidden demand) are ignored. In addition, past model construction has often emphasized simulation prediction ability, paid attention to the model parameter check and quantitative simulation result description, and has lacked comprehensive and in-depth

elaboration of the model framework. As a result, the simulation process has become a black box due to its extreme complexity and insufficient explanation of the failures of past control policies. More importantly, previous models have ignored the subjective initiative of the government to formulate relevant policies to address specific problems; that is, the effect of traffic congestion on the promotion of anti-congestion policies. Therefore, previous models lacked the inclusiveness of emerging congestion control policies and faced difficulty in evaluating specific policies in a targeted way.

Based on these observations, We propose the following hypothesis:

Hypothesis: The "Sparse Block Collocation" Policy was probably the most sustainable and fundamental congestion control measure.

Then, this study uses an SD regression to improve the basic goal of the mental model and highlight the analysis function of the SD model to deduce the influence of policy implementation. The SD method is used to analyze current traffic congestion control measures in China and determine which control measures are the most effective and sustainable. By depicting the causal feedback diagram of the congestion mechanism (also known as the system flow diagram, which refers to the organic combination of the basic variables of SD and its symbols, the internal structure of the system, and its feedback mechanism for qualitative analysis), this study answers the following question: Which popular transportation policies are truly effective and sustainable? And the main aim of this work is to provide consult for the government to put forward reasonable and effective measures to control traffic congestion.

2. Materials and Methods

2.1. SD Method. SD was founded in 1956 by Jay Forrester and is a method used in the quantitative research of complex social and economic systems based on feedback control theory and computer simulation technology [14]. The SD method is realized by establishing a feedback loop (also known as a causal feedback loop) and setting various variables and equations. A feedback loop refers to a closed causal chain sequence composed of different elements or variables in a system. In 1988, McLaughlin published the book Application of Systems Approach to Urban and Regional Planning in which he asserted that "cities can be regarded as dynamic systems that are constantly changing under the influence of many factors, so the planning of cities with dynamic systems must also be dynamic planning with constant changes" [15] (p.105).

Forrester published Urban Dynamics in 1969; however, SD was only introduced in China in 1980, with the first national academic exchange conference on the topic held in 1986. Since then, many institutions have both conducted research on and considered the applications of this concept [16]. For example, in the field of urban transportation, studies have examined private car ownership in China. Li and Yaping studied urban traffic development using the

dynamics model of the private car consumption system in Beijing. It was based on system structure and causal feedback analysis; and it discussed the restraining effect of the development of public transportation, especially rapid rail transit, and private cars development; and provided a decision-making basis for the formulation of an urban transportation development policy [17]. The models and methods mentioned above are similar and develop gradually, providing inspiration and reference for this study.

2.2. Building the Formation Mechanism Model of Traffic Congestion.

The most important aspect of drawing a causal feedback map of congestion mechanisms is to understand that there are two feedback loops in the system: positive and negative. Each feedback loop is composed of more than two variables with a clear causal relationship between the variables, and there are two kinds of positive (+) and negative (-) causal relationships. A positive feedback loop refers to when the number of negative causalities in the feedback loop is even. A negative feedback loop is one in which the number of positive causalities is odd. Positive feedback loops are characterized by self-enhancement and self-amplification [18]. Typical examples are the chicken–egg and egg–chicken (①⟶②⟶①, Figure 1, a). Negative feedback loops are self-controlled and self-weakening. A typical example is an effort made to achieve a certain goal. For example, students can determine their future degree of study efforts according to the gap between their test scores and target scores to ensure that their academic performance will be closer to their target through their constant effort (③⟶④⟶⑤, Figure 1, b). In this study, the comprehensive judgment of traffic congestion is the result of a cycle caused by urban construction and traffic environment deterioration under the guidance of a self-enhancement mechanism.

2.3. Method of Analysis

2.3.1. Operational causal chain

Definition 1. If the simulation equation of the causal chain from the variable X(t) to the dependent variable Y(t) is the function Y(t) = f(X(t)), wherein f is the operation symbol of the function, then the causal chain $X(t) \longrightarrow^f Y(t)$ that gives the simulation operation symbol is called the operation causal chain.

Definition 2. If we run the causal chain $X(T + k * DT) \longrightarrow^f Y(T + k * DT)$, it is called the causal chain of simultaneous operation of simulation step size. If the operation causal chain is $X(T + (k-1) * DT) \longrightarrow^f Y(T + k * DT)$, it is called the causal chain of simulation step length heteroscale operation.

Flow level variable simulation step length multi-layer operation causal chain. Since the flow bit $L_i(t)$ has the following function relationship:

$$L_i(T + k * DT) = L_i(T + (k-1) * DT) +$$
$$DT * R_{i1}(T + (k-1) * DT) \qquad (1)$$
$$- DT * R_{i2}(T + (k-1) * DT).$$

Then T is the starting time of simulation, and the left side of the function of the flow bit variable $L_i(t)$ is the simulation step length of $t = T + k * DT$, and the right side is the simulation step time of $t = T + (k-1) * DT$, with a difference of one simulation step DT.

The value of $L_i(T + k * DT)$ on the left side of the flow potential variable is the algebraic sum of the three terms on the right side of $L_i(T + (k-1) * DT)$, the inflow rate $R_{i1}(T + (k-1) * DT)$, and the outflow rate $R_{i2}(T + (k-1) * DT)$. There are three corresponding causal chains from right to left:

$$L_i(T + (k-1) * DT) \longrightarrow L_i(T + k * DT),$$
$$R_{i1}(T + (k-1) * DT) \longrightarrow L_i(T + k * DT), \qquad (2)$$
$$R_{i2}(T + (k-1) * DT) \longrightarrow L_i(T + k * DT).$$

The three chains from right to left of the flow potential variable are endowed with different operations to form the algebraic sum of the three terms and the functional expression:

① $L_i(T + (k-1) * DT) \longrightarrow L_i(T + k * DT)$, which is the addition of the three algebraic sums, the "+" operations, and the causal chain with the "+" symbol:

$$L_i(T + (k-1) * DT) \xrightarrow{+} L_i(T + k * DT). \qquad (3)$$

In this study, "+" in the causal chain represents the "plus" operation symbol, but this does not mean that the causal chain polarity is positive. The "-" is in the middle of the causal chain and represents the "minus" sign in the operation. The same applies to the operations below.

② $R_{i1}(T + (k-1) * DT) \longrightarrow L_i(T + k * DT)$ is the addition to the algebraic sum of three terms and is also the plus term. However, there is an operation term of $*DT$ again, so there is a causal chain with a "+ " and " $*DT$" double operation symbol. It is marked as follows:

$$R_{i1}(T + (k-1) * DT) \xrightarrow{+, *DT} L_i(T + k * DT). \qquad (4)$$

③ $R_{i2}(T + (k-1) * DT) \longrightarrow L_i(T + k * DT)$ is the subtraction item in the three-term algebraic sum and is the "-" operation item and also the operation item with $*DT$, so there is a causal chain with a "-" and " $*DT$" double operation symbol.

$$R_{i2}(T + (k-1) * DT) \xrightarrow{-, *DT} L_i(T + k * DT). \qquad (5)$$

The above three operation causal chains reveal that the causal chain of the flow potential variable $L_i(T + k * DT)$ is composed of the following three causal chain diagrams of the different layer operations. Figure 2 shows that the vertex of the flow potential variable in the calculation causal chain graph is the frame vertex in the hold flow graph.
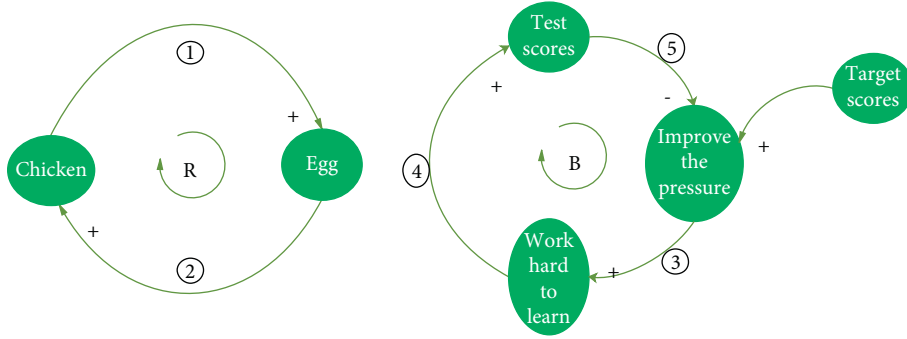
FIGURE 1: Example of reinforcing loop (a) and balancing loop (b).
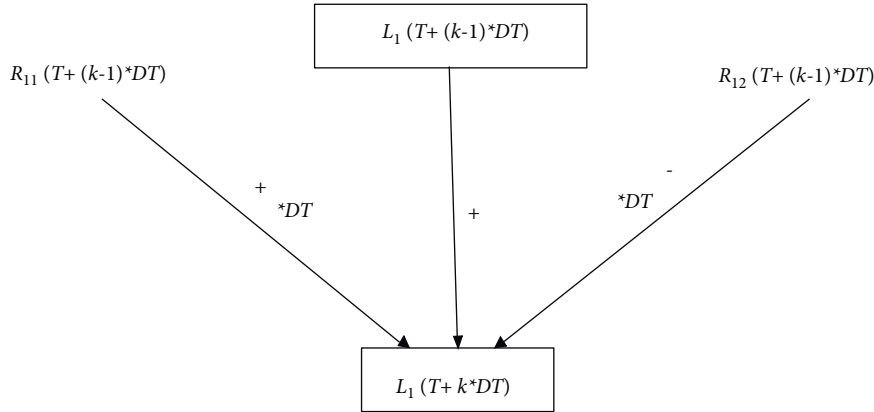
FIGURE 2: Flow level variable $L_i(T + k * DT)$ of the different layer operation causal chain diagram.

Proposition 1. *The causal chain of flow potential variables operation in the system dynamics flow diagram is all 3-layer operation causal chain of flow potential variables $L_i(T + k * DT)$. It is the combination of the causal chain of the flow potential variables in the flow diagram and the simulation equation:* $L_i(T + k * DT) = L_i(T + (k-1) * DT) + DT * R_{i1} (T + (k-1) * DT) - DT * R_{i2}(T + (k-1) * DT)$.

Intuitively revealing the causal chain of different level operation, the value of the flow level variable $L_i(T + k * DT)$ is its last moment $L_i(T + (k-1) * DT)$ value, plus the inflow rate times the step size $DT * R_{i1}(T + (k-1) * DT)$ value, and the minus outflow rate times the step size $DT * R_{i2}(T + (k-1) * DT)$ value. Combining the causal chain with the simulation equation, the internal relationship between the causal chain and the simulation equation can be established, and the intuitiveness of the model can be enhanced.

Flow rate variable simulation step length time simultaneous operation causal chain. After the simulation calculation of flow level variable of the system dynamics model, the simulation calculation of flow rate variable is carried out, and the simulation calculation of flow rate variable is the same as the simulation step length calculation. For example, the quotient equation of outflow rate $R_{i2}(t)$: $R_{i2}(t) = L_i(t)/D_1$. The independent variables on both sides of the equation are t, and in the flow diagram, there exists a causal chain from the flow potential variable to the outflow rate:

$$L_i(t) \longrightarrow R_{i2}(t). \tag{6}$$

The operator of this quotient equation is $1/D_1$, at the $t = T + k * DT$ simulation step length, the causal chain of its parallel operation is:

$$L_i(T + k * DT) \overset{1}{\longrightarrow} /D_1 \; R_{i2}(T + k * DT). \tag{7}$$

The flow rate causal chain is the combination of the flow level variable to the flow rate causal chain and its simulation equation.

According to the definition, in the flow diagram of system dynamics, flow rate to flow rate, flow rate to auxiliary variable, auxiliary variable to flow rate, and auxiliary variable to auxiliary variable are all causal chains of the same layer operation.

Proposition 2. *In the dynamic flow diagram of the system, the causal chain of the most operation of flow level change is the causal chain of the simulation step length and time different layer operation, and the causal chain of the other variable operation is the causal chain of the simulation step length and time the same layer operation.*

Proposition 3. *System dynamics any simulation step length time $t = T + k * DT$. First of all, according to the causal chain of each flow potential variable with 3 different levels of operation, the $T + (k-1) * DT$ correlation multiplier of the*

previous simulation step is used to calculate the causal chain of each flow potential variable, and the values of all flow potential variables $t = T + k * D$ time are calculated. Then, do the same level operation for the other variables, the causal chain calculation, and calculate the values of $t = T + k * D$ time for all the other variables.

The above three propositions are the default requirements of the system dynamics simulation program.

2.3.2. Operation process diagram

Definition of operation process diagram. Let $G = (V(t), X(t))$ be an ordered binary, $V(t)$ is the vertex set, $V(t) = \{V_1(t), V_2(t), \ldots, V_{n-1}(t), V_n(t)\}$; $X(t)$ is the directed arc set, an arc is an ordered pair of $(V_i(t), V_j(t))$ composed of different elements of $V_i(t) \cdot V_j(t)$ of the vertex set $V(t)$. Then $G = (V(t), X(t))$ is called a directed graph.

Definition 3. directed graph $G = (V(t), X(t))$, t is the simulation step length time $T + K * DT$, $V(T + K * DT)$ is the vertex set of data process, arc set $X(T + K * DT)$ is the set of operation causal chain of assignment operation symbol, an ordered binary set of vertex sets of data procedures and causal chain sets of operations is called an operation procedure graph, recorded as $G = (V(T + K * DT), X(T + K * DT))$.

Definition 4. The computation process diagram of the simulation model is an ordered binary of the vertex set of the data process and the causal chain set of the computation.

The establishment steps of operation process diagram

Step 1. Based on the simulation table or some known conditions, the causal chain of flow graph variable is combined with the simulation equation, and the operational causal chain graph without calculating the vertex value is established layer by layer.

Step 2. The variable data process of calculating the uncalculated vertex values of each layer of the causal chain graph, showing the calculation process of variable vertex values.

3. Results and Discussion

3.1. The Logic of Traditional "Road Construction and Congestion Control" Policies. From the perspective of SD, this section describes two disadvantages of traditional road construction and congestion control strategy: the release of short-term hidden demand and the generation of medium - and long-term demand. To some extent, this opens the analysis and evaluation of several congestion control measures in the following discussion.

The direct cause of traffic congestion is an imbalance between the demand for car travel and the supply of road facilities. In recent decades, China's urbanization process has accelerated. An increasing number of people have entered cities, and living standards have been continuously improved. Private cars have gradually transformed from high-

end luxury goods to mass-consumption goods. Under the combined force of an increase in the population base and a rise in the private car ownership rate, the demand for car travel has increased, thereby rendering transportation infrastructure in many cities insufficient to support this shift, resulting in added traffic congestion [19].

Regarding public complaints, the government's pressure on congestion control has increased, and road construction has long been the dominant solution. Road widening, viaduct construction, expressway construction around the city, and other unimpeded projects are common in various cities in China. In addition, driven by the desire for image projects and urban expansion, Chinese city governments invest billions in road planning, construction, and management annually. This infrastructure is given to the car group free of charge, along with many modern traffic managers and management methods. Furthermore, 70% to 80% of traffic researchers' efforts serve car users [20].

From the perspective of SD, the causality of ①⟶②⟶③⟶④ groups form a closed negative feedback loop, thereby indicating that road construction can alleviate local traffic congestion in the short term (Figure 3). If traffic was simply a self-debilitating mechanism as many government officials and technologists imagine it to be, then the problem would be solved. However, this is not the case because there is a series of self-reinforcing mechanisms that play a more dominant role.

3.1.1. Post-Road Effect 1: The Release of Short-Term Hidden Demand. When the government was occupied by the positivity of road construction and ribbon cutting, urban congestion was alleviated immediately. However, in the short term, the appeal of car travel was rapidly improved, which directly encouraged people to drive both more frequently and for longer distances, thereby leading to the increase of car travel and the subsequent congestion of new roads (negative feedback loop ①⟶②⟶③⟶④, Figure 4). Conversely, the appeal of car travel became the most important factor in encouraging citizens to buy private cars, other than an increase in income, which also led to an increase in car travel volume and the deterioration of road congestion (negative feedback loop ①⟶⑤⟶⑥⟶⑦⟶④, Figure 4). Figure 4 shows that there are two negative feedback loops in the process of eliminating the congestion control effect caused by road construction, which indicates that the suppressed hidden demands of the original congestion period are released and transformed into explicit demands, thus aggravating the road burden.

However, the four-stage traffic model of "trip generation - trip distribution - mode division - road network allocation," widely adopted by transportation professionals, cannot accurately simulate the hidden demands, which leads to the conclusion that more roads can solve congestion. Although traffic modelers assume future increases in travel demand in traffic forecasting scenarios, these demands are often considered the product of economic growth and car ownership. In this model, expanding the supply of road
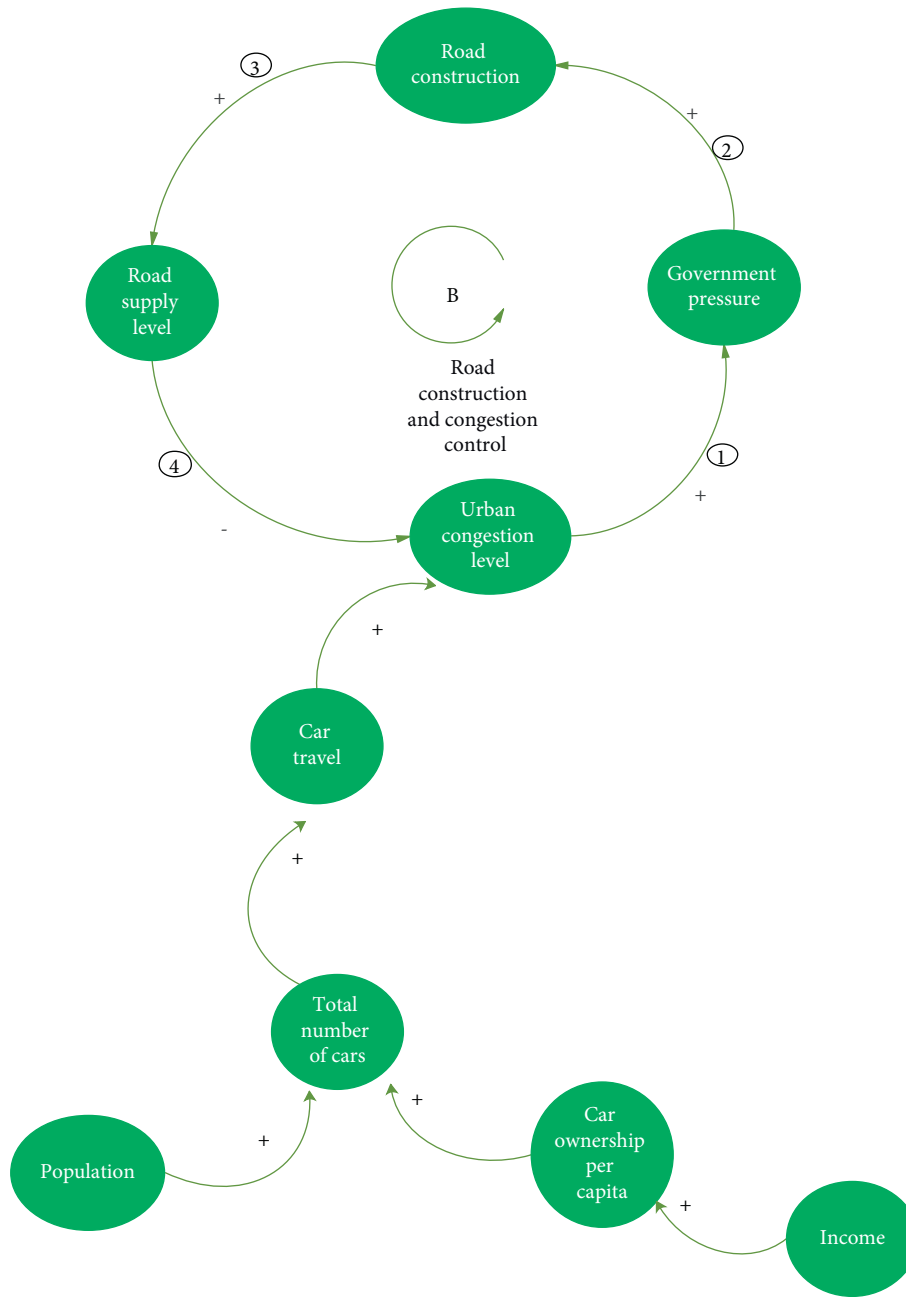
FIGURE 3: Conventional logic of building more roads to relieve congestion.

infrastructure does not affect the forecast of future transport demand growth [21]. In other words, the importance of latent demand has been ignored.

3.1.2. Post-Road Effect 2: Generation of Medium- and Long-Term Derivative Demand.

In the medium- and long-term perspectives, road construction will also lead to a cycle of "congestion, road construction, congestion, and road construction" (Figure 5) because the government builds roads on the outer edge of the city to enclose land. This construction usually adopts a car-led development model, thereby resulting in a large amount of induced demand.

Three positive feedback loops account for the rapid deterioration of urban traffic congestion in China, as follows.

First, the imbalance between jobs and housing in new areas led to the emergence of several sleeping cities around central cities (such as Huilongguan and Tiantongyuan in Beijing). The migration of residents has greatly increased the daily commuting distance, which causes congestion. However, the government continues to build roads to control congestion, which causes a vicious cycle (positive feedback loop ①⟶②⟶③⟶④⟶⑤⟶⑥⟶⑦⟶⑧⟶①, Figure 5).

Second, large new urban areas accommodate more urban populations, thereby objectively increasing the total
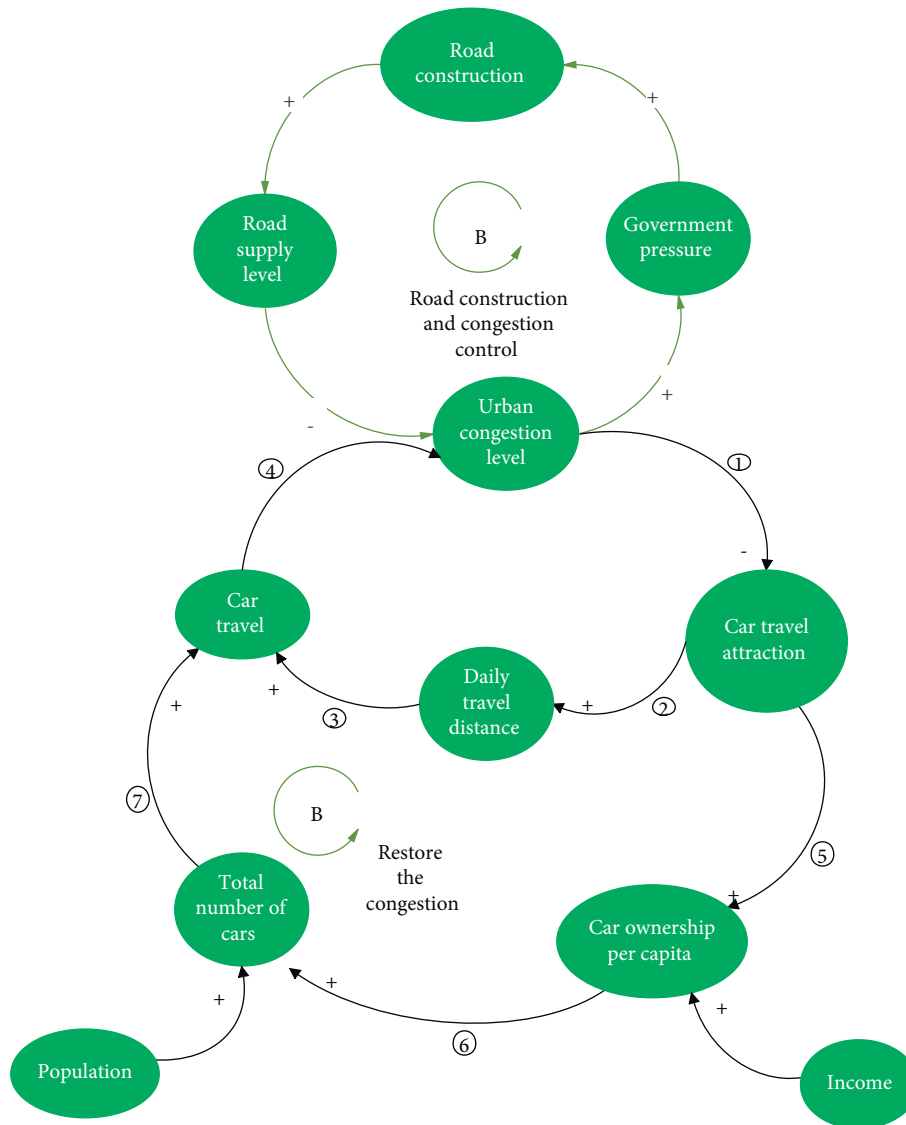
FIGURE 4: Short-term release of latent demand after road building.

population base of urban travel and the congestion of the traffic system. Highstreets and large road networks lead to low permeability of bus systems, and meeting client satisfaction regarding speed and efficiency is a common problem. Regarding the slow traffic environment, the dredging and construction of the main roads block the microcirculation of the slow traffic network, resulting in the need for a long-distance detour to avoid crossing the road. However, intersections of wider roads form larger intersections, which causes excessively long waiting times at red lights, traffic disorder at intersections, and the need to climb the overpass under the underground passage; all of which causes inconvenience. Additionally, the car parking management in the new districts is loose, and the phenomenon of bicycle lanes and sidewalks being occupied by parked cars is prominent. Furthermore, because of the lack of bicycle parking facilities, parking is difficult, and it is easy to lose bicycles at destinations due to poor parking management. Subsequently, bicycles have become the second transportation choice, even for short- and medium-distance travel, thereby augmenting the reliance on car travel (positive feedback loop ① ⟶ ⑫ ⟶ ⑬ ⟶ ⑭ ⟶ ④ ⟶ ⑤ ⟶ ⑥ ⟶ ⑦ ⟶ ⑧ ⟶ ①, Figure 5).

Similarly, the third loop involves the deterioration of public transportation and slow-moving environments and has also led to a surge in the car ownership rate in new districts. Many people regard staying in large districts and buying private cars as binding lifestyle transition nodes (positive feedback loop ① ⟶ ⑫ ⟶ ⑬ ⟶ ⑮ ⟶ ⑯ ⟶ ⑪ ⟶ ⑤ ⟶ ⑥ ⟶ ⑦ ⟶ ⑧ ⟶ ①, Figure 5).

The three positive feedback loops outlined above are the main reasons for the rapid deterioration of urban traffic congestion in China. Taking Beijing as an example, in the last 10 years, the functions of Beijing have been concentrated in the central urban area, thereby resulting in a large increase in the proportion of long-distance travel that is difficult to complete on bicycles. According to a survey by the Beijing Transportation Development Research Center, the
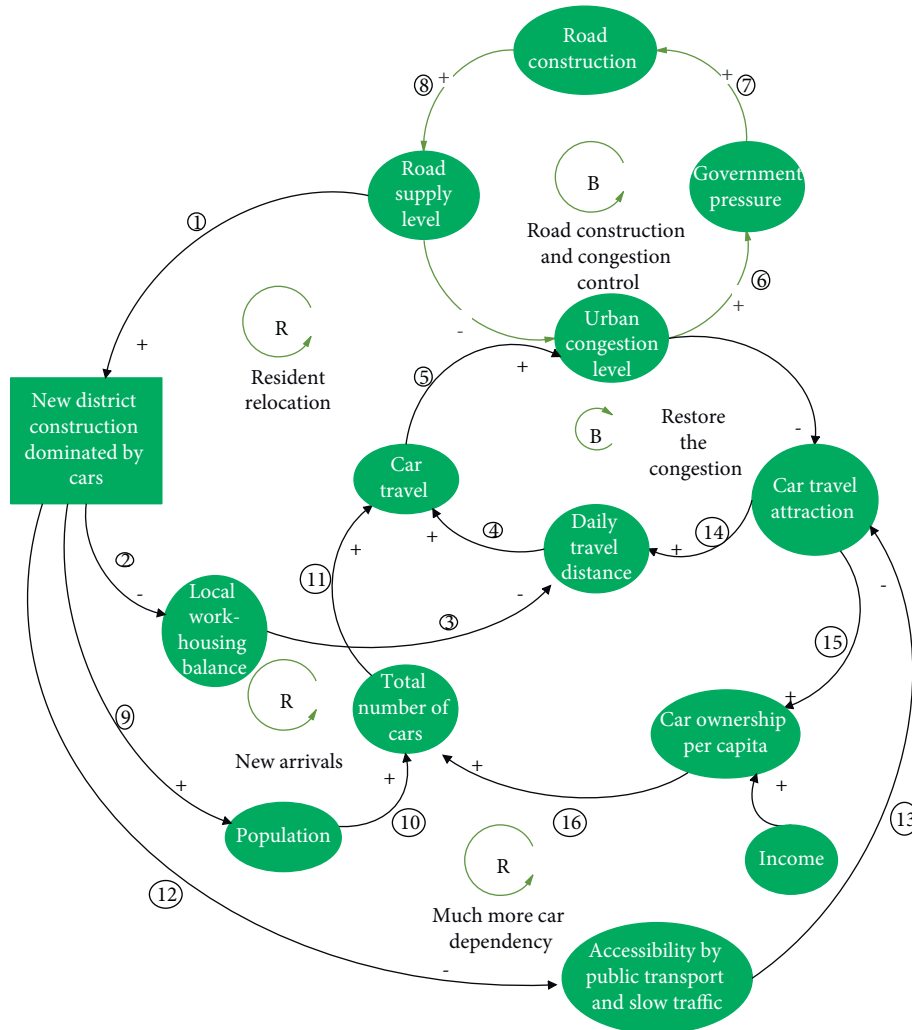
FIGURE 5: Medium- and long-term generation of induced demand after road building.

proportion of bicycle trips in Beijing dropped from 30.3% in 2005 to 17.9% in the first half of 2019 [22]. Among the car trips, the proportion of short-distance trips within 5 km was more than 40% [23]. The reason for driving is because the environment for walking and cycling is deteriorating, which is closely related to the new urban form that is dominated by large blocks and wide roads.

The above analysis shows the mechanism of the release of implicit demand for car travel and the generation of the derivative demand in the post-road construction period. It can be predicted that when per capita income reaches a point where it is no longer a major barrier to household car consumption (for a significant portion of households), by allowing the ability to choose to travel by motor vehicle, traffic congestion will follow a barrel effect with slow bus accessibility as a shortcoming.

3.2. Analysis of Several Congestion Control Policies. According to the framework in the above section, the congestion control policies can be divided into five categories: similar road construction, rigid intervention, supply

adjustment, demand guidance, and sparse block collocation. Using the framework established in the previous section, we next analyze the congestion control policies.

3.2.1. Similar Road Construction Policies. Such policies include increasing the supply of parking spaces, improving the service level of intersections (building flyovers and pedestrian overpasses), enhancing the efficiency of road networks (using intelligent transportation systems and creating green wave corridors for vehicles), and so on [24]. The goal of these policies is to increase the space supply of motor vehicles or expand the capacity of road traffic by summarizing them as "similar road construction and congestion control" policies. Similar to road construction, closing the negative feedback loop indicates that the road supply level can indeed increase in the short term. However, it will soon be consumed by the released hidden demand and the newly-derived demand, thereby exacerbating traffic congestion further (Figure 6).

3.2.2. Rigid Intervention Policies. Rigid intervention policy measures include population control, license plate lottery,

FIGURE 6: Consequences of adopting similar road construction policies.

and car license plate restrictions, which are characterized by direct and forcible interventions in the growth of key elements within the traffic system (such as population and cars or car travel volume). Although these measures will have an immediate effect in the short term, they are often conducted at the expense of higher personal and social costs and are suspected of tending toward groups with vested interests. For example, after the restrictions on these key elements were implemented, the need to shift to the public transport system made the already crowded situation worse [25]. Meanwhile, such policies only temporarily transform part of the explicit demand into implicit demand, which is likely to be re-released when the policy is relaxed, thereby resulting in the rapid recovery of the system and the aggravation of congestion.

Many drivers and traffic police groups attribute traffic congestion to the low civilization quality of pedestrians and cyclists (such as running red lights, jaywalking, etc.), and believe that traffic civilization construction can be strengthened to alleviate traffic congestion [26]. However, without discussing the effectiveness and lag of the construction of traffic civilization, this can be attributed to the following reasoning. If from tomorrow, citizens begin to strictly observe the traffic order (such as not running red lights), then the traffic order may be greatly improved in a short time, thereby significantly improving the driving efficiency of motor vehicles. However, this improvement can only be made possible by inconveniencing pedestrian or cycle travel (such as the introduction of a long wait at traffic signals),

which, in the long term, will encourage people to drive, once they are able to buy a car, thus adding to the congestion[27].

3.2.3. Supply Adjustment Policies.

In contrast to the "rigid intervention type" policies, supply adjustment policies start from the "supply side," prioritize slow or public transportation and attempt to introduce a new negative feedback loop to achieve a self-reduction mechanism of congestion (Figure 7). By reducing unnecessary vehicle space (such as avoiding excessively wide vehicle lanes, reducing the turning radius of intersections, and adding safety islands), the crossing distance for slow pedestrians can be shortened, the speed of vehicles can be reduced, and the safety and convenience of walking and cycling can be improved. This result will indirectly improve the accessibility of the public transportation system and reduce the dependence on cars, cars' travel distances, and the rate of private car ownership, thereby reducing the number of car trips and alleviating traffic congestion (negative feedback loop ①⟶②⟶③⟶④⟶⑤⟶⑥⟶⑦, Figure 7; negative feedback loop ①⟶②⟶③⟶④⟶⑧⟶⑨⟶⑩⟶⑦, Figure 7).

The slimming of roads may seem to reduce the amount of space available for vehicles, but it does not. So, by optimizing the organization of roads and effectively reducing the interference of mixed traffic, it is possible to achieve a win-win situation for slow traffic and motor vehicles. In recent years, some foreign cities (such as New York, Copenhagen, Seoul, etc.) have boldly transformed urban motorways into slow lanes or bus lanes. As a result, they have not only improved the vitality of the city and the quality of public spaces but have also eased traffic congestion, which has been welcomed by citizens. For example, the renovation project of Broadway Avenue in New York transformed a congested road into a special strip for pedestrians to walk and rest (Figure 8). According to a recent evaluation study, the speed of vehicles in the area increased by nearly 17% after the project was implemented [28].

Another example is the Bus Rapid Transit (BRT) test line in Guangzhou, which was built in the most congested area where buses and social vehicles had serious interaction before its reconstruction. By turning the motor lane in the middle of the road into a bus lane, the running speeds of both public and social vehicles were improved (Figure 9). After the BRT corridor was put into operation in February 2010, the average speed of social vehicles along Zhongshan Avenue increased by 28% compared to before [29].

3.2.4. Demand-Oriented Policies.

From the perspective of demand management, the policies of economic leverage to increase cars' travel costs have increased parking fees, levied congestion charges, and strengthened the restrictions on unregulated parking. These measures can effectively restrain the uncontrolled use of cars during rush hour (negative feedback loop ①⟶②⟶③⟶④⟶⑤⟶⑥, Figure 10) and can simultaneously slow down the popularity of cars, reduce the number of car trips, and alleviate the degree of congestion (negative feedback loop ①⟶②⟶③⟶⑦⟶⑧⟶⑨⟶⑥, Figure 10).

Another benefit of increasing parking fees and congestion pricing is the amount of economic revenue that the city government can generate, which can then be used to subsidize the construction or improvement of the efficiency of the public transport system [30]. It can also improve the slow travel environment, increase the attractiveness of short- and medium-distance slow travel, and solve the last mile problem for long-distance bus travel. All these changes will reduce citizens' reliance on cars for travel, the distance traveled by cars, the rate of private car ownership, and car trips, thereby alleviating congestion (negative feedback loop ①⟶②⟶⑩⟶ 11 ⟶12 ⟶④⟶⑤⟶⑥, Figure 10; negative feedback loop ①⟶②⟶⑩⟶ 11 ⟶12 ⟶⑦⟶⑧⟶⑨⟶⑥, Figure 10). Similar measures have been adopted in London since February 2003, where part of the revenue from traffic congestion fees have been used to improve public transport in the region. As a result, not only did 20% fewer cars enter the city center every day, but buses were also 25% faster than before, thus reducing the delay time of motor vehicles in the congested charging areas during peak hours by 30% (Figure 10) [31].

3.2.5. Sparse Block Collocation Policies.

With the aim of increasing the construction of public transport systems and adopting the transit-orientated development (TOD) model in the construction of new areas around bus stations, the policies are to: (1) design walkable streets and pedestrian-scale blocks to enhance pedestrian traffic; (2) incorporate pedestrian safety and convenience requirements into architectural design; (3) reduce the demand for motor vehicles by creating bicycle-friendly road networks; (4) increase the use of public transport by building public transport-oriented streets and communities; (5) advocate mixed land-use patterns to disperse public travel destinations; and (6) establish public green spaces and services within walking distance of each other [32].

The impact of such policies can be explained by the following series of negative feedback loops (Figure 11).

Mixed land use enables more balanced employment and housing, shortens commuting distances, and reduces traffic demand. Simultaneously, it is possible to use bicycles and walk for short and medium distances, which increases the proportion of green travel and reduces car travel (negative feedback loop ①⟶②⟶③⟶④⟶⑤⟶⑥, Figure 11).

Increasing the construction of public transportation facilities and attractions of public transportation can reduce the dependence on cars. Owing to the distribution of attraction points along the bus corridor, pendulum traffic flow can be avoided. The reduction of the one-way distance per person can improve the bus carrying turnover rate and avoid one-way traffic congestion during peak hours. The development of small blocks can reduce bottlenecks along major and secondary roads and large intersections, thereby enhancing the permeability of public transportation services and solving the problem of the last mile by improving the
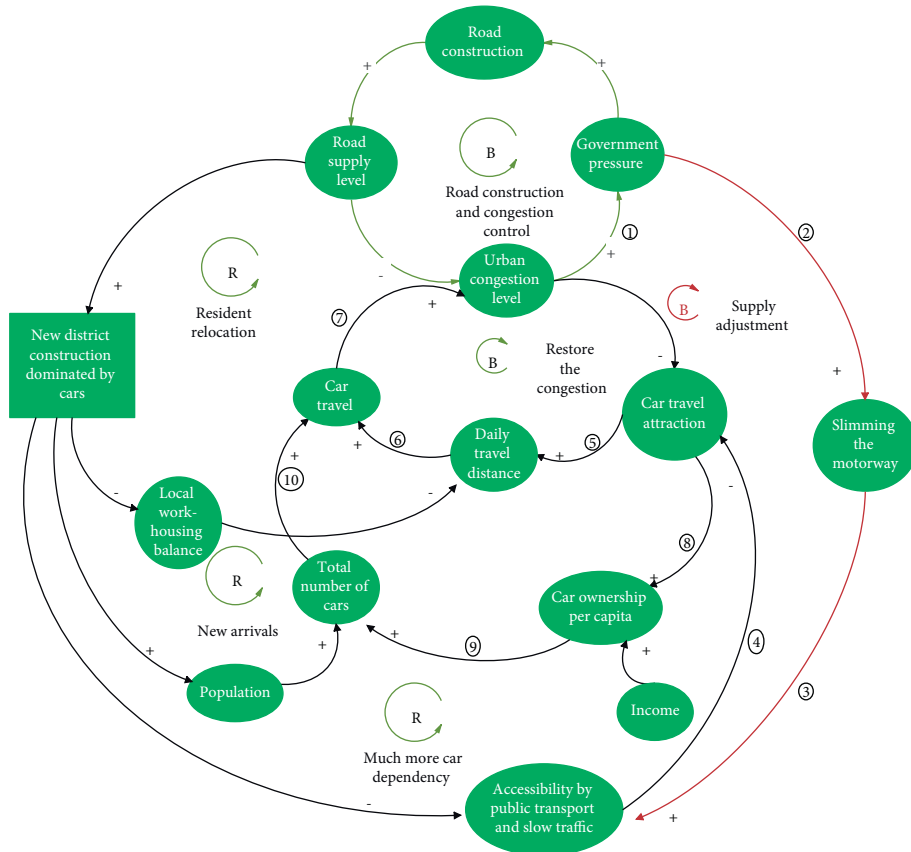
Figure 7: Consequences of adopting adjusting supply policies.

Figure 8: Before/after comparison of the Broadway renovation project in New York.

slow traffic environment. The impact of these aspects can reduce the attractiveness of car travel by lowering the private car ownership rate, car use intensity, and a number of car trips, thereby alleviating traffic congestion (negative feedback loop ①⟶②⟶⑦⟶⑧⟶⑨⟶⑤⟶⑥,

Figure 11; negative feedback loop ①⟶②⟶⑦⟶⑧⟶⑩⟶ 11 ⟶12 ⟶⑥, Figure 11).

If the number of parking spaces is reduced and the private car ownership rate of residents in TOD areas can be

Figure 9: Before/after comparison of the Zhongshan Avenue renovation in Guangzhou.
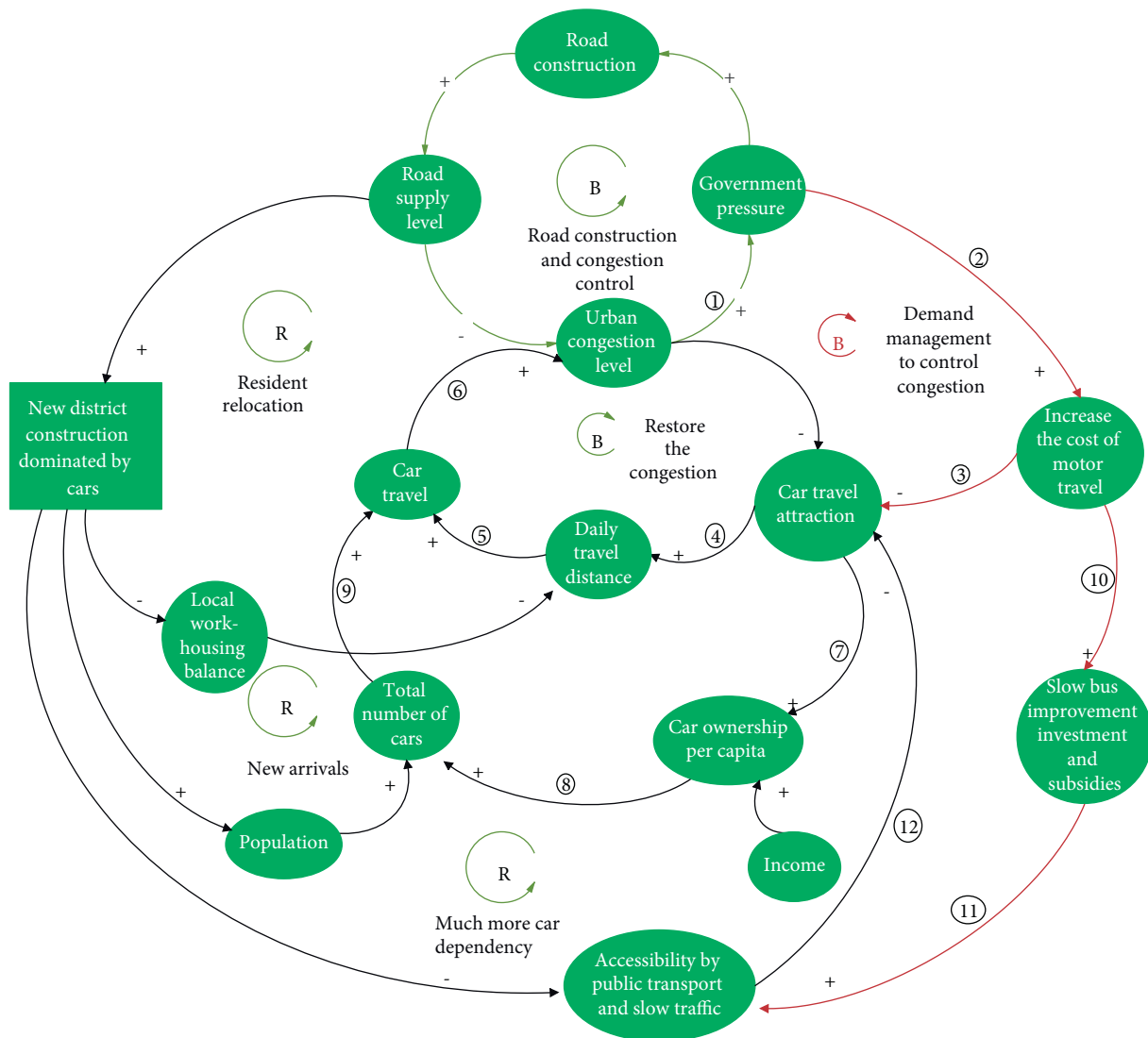
Figure 10: Consequences of adopting demand-oriented policies.

effectively controlled to a low level, the number of car trips can be reduced and congestion alleviated (negative feedback loop ①⟶②⟶ 13 ⟶14 ⟶11 ⟶12 ⟶⑥, Figure 11).

Compared with previous types of congestion control policies, public transit-guided development and sparse block collocation policies can establish multiple and effective congestion mitigation mechanisms from "open sources" (attracting car travel to other modes of transportation) to "throttling" (reducing basic travel volume), thereby creating more comprehensive and further-reaching traffic control.
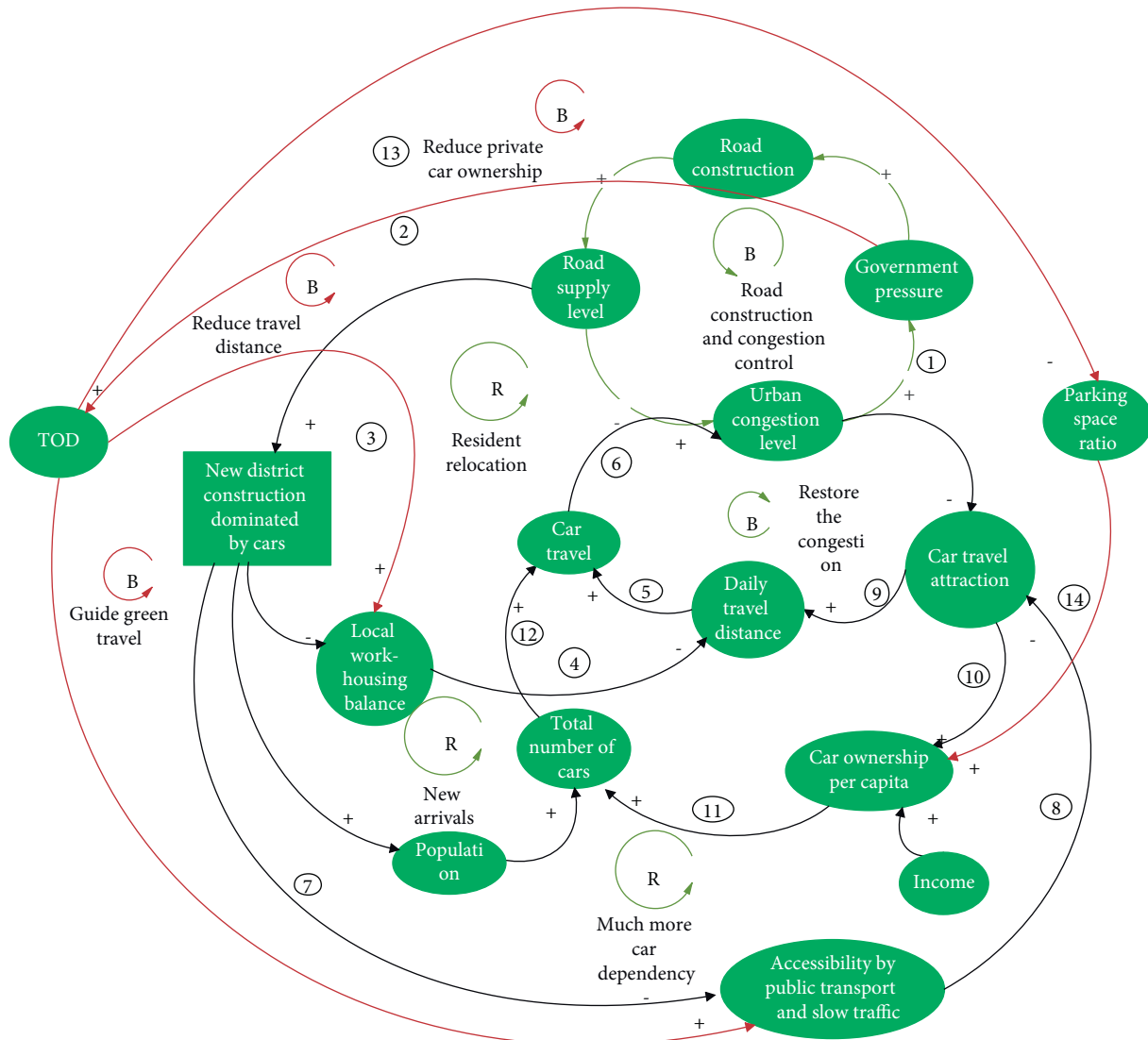
Figure 11: Consequences of adopting TOD policies on the traffic demand and supply simultaneously.

4. Conclusions

This study builds an urban congestion mechanism model based on SD theory. The analysis shows that although traditional road construction and control congestion methods can alleviate congestion in a short period of time, the huge hidden demand and the development of car-oriented new districts caused by road construction will be prevalent in the future. Under the dual influence of the derived demand, traffic congestion will quickly recover and intensify. Unfortunately, traffic congestion cannot be solved by traffic engineers alone, as traditional traffic model predictions cannot simulate these interactions.

Faced with the pressure to control congestion, the government has adopted five policy types: similar road construction, rigid intervention, supply adjustment, demand guidance, and sparse block collocation. This study introduces the SD model to analyze these policies and shows that

similar road construction policies continue to misunderstand traditional road building to prevent traffic congestion. Although rigid intervention policies have immediate effects in the short term, they will be at the expense of personal convenience and social cost, and the suppressed hidden demand will eventually be released, making it unsustainable overall. The supply-adjusted policies and demand guidance policies are relatively more effective, not only for short-term results but also for long-term sustainability. Meanwhile, the sparse block collocation policies are the most sustainable and fundamental congestion control measure, with public transport-oriented land use and development as its core. It forms multiple negative feedback loops to relieve traffic congestion from both the supply and demand aspects. Although the land development policy is often regarded as a long-term policy, it is both urgent and necessary to take effective measures to curb its use, and instead, implement the concept of public transport-oriented development in urban

planning policy formulation in view of the rapid development of new areas dominated by cars in China at present.

The fundamental idea of congestion control should be people-oriented, not car-oriented. In the debate between mobility and accessibility, the government has always pursued the mobility of the road, which practices the neglect of essentials. Some officials have argued that the fundamental goal of tackling urban traffic congestion is to guarantee the basic right of residents to travel conveniently, including the freedom to travel by car [33]. In many cases, urban decision-makers often verbally advertise "slow travel first and people-oriented" congestion control measures; however, the ultimate goal is to increase the speed of motor vehicles. This fact is based on the premise of "not harming the interests of car travel" as opposed to the premise of "slow driving and people-oriented" measures, which would be accused of increasing blockages [34]. This logic ignores the harsh reality that given the number of people and high densities of China's cities, the rapid growth of China's economy, and the shortage of oil resources, there is probably no solution that will ultimately satisfy the right of every Chinese residents to drive.

It should be recognized that a small number of people in cities travel purely for the enjoyment of transportation and that most people travel to attain social, recreational, educational, employment-related, and other opportunities and benefits from destinations [35]. Therefore, the government's focus should not be on keeping roads clear (increasing mobility), but on improving accessibility and providing the easiest and cheapest way for citizens to travel. Scientific urban planning and appropriate public policies can achieve twice the result with half the effort. This study found that the sparse block collocation policy is the most sustainable strategy, but did not analyze the strategy's implementation details. Future research should start with the concrete implementation details of the sparse block collocation policy.

Data Availability

The data used to support the findings of this study are available from website of National Bureau of Statistics.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] Beijing Times and Survey says over 80% of car owners suffer from "traffic jam dysphoria, "Survey says over 80% of car owners suffer from "traffic jam dysphoria," 2011, http://www.ckxxbao.com/baozhi/jinghuashibao/ (accessed on 23 April 2011.

[2] G. Minghua and Z. Zhenghe, "The solution to the problem of traffic congestion in big cities," *City Prob*, vol. 10, pp. 75–78, 2009.

[3] W. Yuanqing, S. Linyan, and S. Chuanjiao, "The system cause and solution of traffic problem in big cities of our country," *City Prob*, vol. 5, pp. 75–80, 2007.

[4] H. Xianglin, J. Changyuan, and G. Hongxia, "Coupled mapping following model and traffic congestion control based on intelligent transportation system," *Journal of Physics*, vol. 8, pp. 4383–4892, 2007.

[5] Shenzhen, "NPC deputies say heavy traffic is not the root cause of congestion," 2011, http://epaper.southcn.com/nfdaily/html/2021-06/18/node_2.htm accessed on.

[6] L. Feng, H. Runlong, and D. Jinhong, "How do megacities regulate population size?" *Population Studies*, vol. 1, pp. 29–43, 2011.

[7] L. Xianteng, S. Qing, and Z. Li, "Contradictions between supply and demand of traffic in big cities and countermeasures for their development -- Taking Nanjing as an example," *City and Planning*, vol. 1, pp. 80–87, 2009.

[8] To Solve Congestion, "We must vigorously build urban public transport system," 2010, http://paper.people.com.cn/rmrb/html/2021-06/18/nbs.D110000renmrb_01.htm.

[9] Beijing Municipal People's Government, *Opinions of Beijing Municipality on Further Promoting the Scientific Development of Transportation in the Capital and Strengthening the Work of Easing Traffic Congestion (Jing Zhengfa [2010] No. 42)*, People's Government of Beijing Municipality, Beijing, 2010.

[10] Guangzhou Municipal People's Government, *Plan for Alleviating Traffic Congestion in Guangzhou Central Urban Area after Asian Games (Draft for Soliciting Opinions)*, Guangzhou Municipal People's Government, Guangzhou, 2011.

[11] W. Jifeng, L. Huapu, and P. Hu, "SD model of urban traffic system and its application," *Transp. Syst. Eng. Inf.*vol. 3, pp. 83–89, 2008.

[12] L. Shuang, *An Empirical Study on the Evolution Mechanism of Urban Traffic Structure Based on System Dynamics*, Beijing Jiaotong University, Beijing, 2009.

[13] Z. Yimei and Z. Yi, "Simulation study on system dynamics model of urban traffic congestion," *Transp. Comput.*vol. 2, pp. 94–101, 2008.

[14] W. Qifan, *System Dynamics*, Tsinghua University Press, Beijing, 1998.

[15] W. Fengwu, *Trans. JB McLaughlin. The Application Of Systematic Method In Urban And Regional Planning*, China Architecture and Building Press, Beijing, 1998.

[16] S. Teng, *System Dynamics Analysis of Planning Policy*, Tsinghua University, Beijing, 2009.

[17] L. Liu and Y. Zhou, "Study on urban traffic development based on the dynamic model of private car consumption system in Beijing," *Urban Dev. Res.*vol. 6, pp. 24–28, 2008.

[18] Z. A. Cong, B. Fl, A. Xl, and D Yuchuan, "Macroscopic modeling and dynamic control of on-street cruising-for-parking of autonomous vehicles in a multi-region urban road network," *Transportation Research Part C: Emerging Technologies*, vol. 128, Article ID 103176, 2021.

[19] M. N. Kamel Boulos, Z. Lu, P. Guerrero, C. Jennett, and A. Steed, "From urban planning and emergency training to Pokémon Go: applications of virtual reality GIS (VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health," *International Journal of Health Geographics*, vol. 16, no. 1, p. 7, 2017.

[20] L. Xiaojiang, "Some thoughts on current urban traffic policy," *Urb. Traff.*vol. 1, pp. 7–11, 2011.

[21] J. D. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill, Boston, 2000.

[22] J. Guo and X. Li, *Beijing Transportation Development Annual Report 2019*, Beijing Transportation Development Research Center, Beijing, 2019.

[23] Q. Yong and S. Mingzheng, "There are several trends in the development of traffic in big cities in China," *Urb. Traff.*vol. 2, pp. 1–6, 2011.

[24] L. Tang, P. Wang, and C. Wang, "Simulation of road transportation energy-saving and emission reduction path based on system dynamics," *Systems Engineering*, vol. 36, p. 6, 2018.

[25] P. Lin, J. Weng, S. Hu, Y. Jing, and B. Yin, "Study on diurnal similarity of individual activity chain of public transport passengers," *Transp. Systems Eng. Inf.*vol. 20, p. 6, 2020.

[26] Z. You, "I analyze the basic causes of urban traffic jam -- from the perspective of management, taking Beijing as an example," *Economics and Management*, vol. 17, 2004.

[27] F Liangbing, L Zhihan, G Gengchen, and S Houbing, "Pheromone based alternative route planning," *Digital communication and networking*, no. 3, pp. 151–158, 2016.

[28] The New York City Department of Transportation, *Green Light for Midtown Evaluation*, The New York City Department of Transportation, NewYork, NY, USA, Report, 2010.

[29] X. Weixiong, "Response measures of tidal passenger flow of bus rapid transit in Guangzhou," *Urb. Traff.*vol. 2, pp. 26–33, 2011.

[30] Z. Jiangping, "Traffic congestion pricing -- the latest international research progress and cases," *City and Planning*, vol. 11, 2010.

[31] Victoria Transport Policy Institute London Congestion Pricing, 2006.

[32] Z. Fengyue, S. Zhenyu, and W. Zhigao, *Low Carbon Transportation: Chenggong New District, Kunming, China Low-Carbon Eco-City Development Report 2011*, China Building Industry Press, Beijing, 2011.

[33] C. Guoqing, "Causes and countermeasures of urban traffic congestion in Beijing," *New Horiz*, vol. 2, pp. 23–25, 2011.

[34] J. Zhu, L. Luo, D. Wang, and D. Ding, *Measurement of Residents' Acceptance of Urban Slow Traffic System: A Case Study of the Pearl River Delta*, , p. 12, Urb. Prob, 2017.

[35] K. Barton, *Transportation Economics*, The Commercial Press, Beijing, 2002.

Research Article

Thermal Simulation Analysis of Internal Control Circuit Board of Steering Gear Box Based on COMSOL Three-Dimensional Simulation Software

Wen Huo [ORCID]

School of Mechano-Electronic Engineering, Xidian University of Technology, Xian 710000, China

Correspondence should be addressed to Wen Huo; huowen@stu.ahu.edu.cn

The steering gear device includes two parts, a steering gear control circuit and a transmission component. The transmission component includes a ball screw and a motor. During the operation of the steering gear, due to the presence of the steering gear ball screw motor and friction, a certain amount of heat will be generated, which will affect the steering gear control circuit in a confined space. At the same time, the steering gear is inevitable in the actual working process, and will experience a high temperature environment, which will increase the temperature of the internal structure of the steering gear, and due to the difference in thermal expansion coefficients between various materials, stress and strain will occur in the structure, which may cause mismatch or even cracks in the system structure, and the steering gear system cannot work normally. It is necessary to analyze the thermal characteristics of the overall steering gear under multiple factors. Based on this, this paper uses COMSOL three-dimensional simulation software to conduct thermal simulation analysis on the shell of the steering gear containing the control circuit board. The temperature distribution and stress-strain response law of the control circuit board in the box, and the influence of different materials and thickness of the box heat insulation layer on the thermal characteristics of the control circuit are discussed, and then a reasonable thickness and material of the heat insulation layer are obtained for the design of the rudder chassis for reference.

1. Introduction

Servo: It is a set of automatic control system composed of DC motor, transmission parts, sensors, and control circuits, which control the output angle of the rotating shaft by sending specific signals. The steering gear generally has a maximum rotation angle. With the progress of the times, electronic packaging devices have increasingly higher requirements for their performance. The substantial improvement in performance has caused a rapid increase in the internal temperature of the packaged devices. Therefore, the thermal, electrical, and mechanical performance requirements for the board-level solder joints in electronic products are also rising. According to Moore's Law proposed by Dr. Moore, every 18 months, the integration

density of semiconductor crystals on active integrated circuits doubles. The control circuit is the control core of the package structure. The dense integration of the control circuit greatly increases the heat generation per unit volume of the heating device such as the chip, and generates more heat. At the same time, the number of solder joints used for the connection of the chip and the circuit board is also increasing. The solder joints in electronic packaging devices are very important in the working circuit [1], and are mostly used for mechanical support and electrical connections. Moreover, studies have shown that the high reliability of the solder joints on the control circuit board usually determines the durability of electronic packaging devices and the working life of the equipment.

2. Related Work

Ye et al. [2] used the GaN (gallium nitride) power amplifier chip as the basic model to analyze the thermal resistance characteristics of its heat transfer path and studied the effects of box material, thermal insulation layer material, and chip size parameters on the heat transfer path. The influence of thermal resistance provides a feasible reference solution for the heat dissipation of power amplifier chips in the future. Once the thermal resistance of the heat transfer path is studied, a feasible reference scheme is given for the heat dissipation of future power amplifier chips. Zhong et al. [3] used high-sensitivity real-time Moiré interferometry to study the interface behavior of flip-chip structures under thermal test conditions. Su et al. [4] taking the chip packaging structure as a model, using infrared thermal imaging technology, studied the change law of the chip thermal stress with the change of current, and it is found that the chip thermal stress changes logarithmically with the change of the working current. At the same time, the relationship between chip thermal stress and operating current was simulated by finite element method to verify its experimental conclusions. The thermal characteristics analysis of the subrack-level electronic packaging structure has achieved certain results under the research of domestic and foreign scholars. Liu et al. [5] studied the structural characteristics, working principle, and heat transfer process of the railway passenger car water purification tank, and established the three-dimensional finite element model of the clean water tank, and its thermal insulation process and heating process are analyzed by transient thermal simulation. The response results are compared with the experimental data to verify the reliability of the model and the research process. Mei et al. [6] used a heavy-duty gearbox as a research model. Aiming at the thermal deformation caused by the thermal load, the thermal characteristics of the gearbox structure were simulated, and the steady-state temperature field distribution of the structure was obtained. The distribution is used as the thermal load, and the response results such as stress, strain, and thermal deformation under the combined action of the thermal load and the structural load are calculated. Qi et al. [7] carried out a three-dimensional finite element modeling of the gear box, combined with the basic theory of heat transfer and finite element analysis theory to conduct a thermal-structural coupling analysis of the gear box, and obtained the temperature field distribution and heat of the gear box. According to the response results, such as deformation, the locations of the maximum stress and the maximum displacement are determined, and the factors affecting the temperature field and the deformation field are analyzed. Xu et al. [8] established a three-dimensional finite element model of a new type of ATCA subrack, and studied the heat dissipation of ATCA chips by changing the structural parameters. On this basis, they studied the influence of air-cooled heat dissipation and fan type on the heat dissipation of the structural box. The finite element simulation of ATCA sub-box structure provides suggestions for the optimal design of the structure. Yoon et al. [9] carried out a three-dimensional finite element modeling of the box structure of a railway bridge, a three-dimensional solid element meshing of the overall structure of the box, and a thermodynamic simulation by applying boundary conditions. Yang et al. [10] used Pro/E three-dimensional software for solid modeling, combined with the basic theory of thermodynamics and finite element analysis methods, and performed thermal-mechanical simulation calculations through finite element software ANSYS, and analyzed the results of its thermal response. The weak links of the box body are removed, and then a theoretical basis is provided for the optimal design of the box body structure. Huo et al. [11] conducted a thermal simulation analysis on the overall structure of the reducer, obtained the overall temperature field distribution of the reducer, and then used the temperature field distribution as the combination of the external thermal load and the static load on the reducer box for thermal-mechanical simulation analysis, Get the stress, strain, and thermal deformation distribution. This paper establishes an overall three-dimensional finite element model that integrates the steering gear control circuit board and the ball screw of the transmission component in a box with a heat insulation layer, calculates the heat source heat of each part of the ball screw of the transmission component, and explores the effect of its heat on the steering gear. The influence of the control circuit, simulating the harsh environment in which the steering gear is in service, placing the two box models with or without heat insulation layer under gradual temperature load for simulation, and comparing and analyzing whether there is heat insulation layer respond in order to explore the effect of the thermal insulation layer on the protection of the internal control circuit of the box. Ming Che et al. [12] used finite element analysis software to simulate the influence of silicon gasket on the heat dissipation of the chip in the package structure. The analysis of the calculation results shows that the better the thermal conductivity of the silicon material gasket, the more heat is transferred, and the purpose of reducing the temperature of the package structure can be achieved. Shuai and Li [13] studied the factors that cause the chip to overheat and the commonly used heat dissipation methods. At the same time, the heating plate is used as a constant heat source to simulate the heating of the computer CPU. The heat dissipation effects of five different types of fans are compared under the various power conditions of the heating plate. Wang and Shen [14] used a multi-chip double-sided PCB board as a model, and carried out finite element simulation by changing the thickness of the PCB board substrate, the size parameters of the chip, and the value of chip power. The analysis found that the above factors all have a large effect on the temperature value.

3. Methods

In the text, the literature on the subject is comprehensively summarized by the method of literature research, and the key literature is refined and summarized, and then the simulation and reproduction of the same model and the same working conditions are carried out through COMSOL, and then the reproduced results are compared with the results in the literature. Carry out comparative analysis, such as simulation cloud map trends, numerical values, and curve changes. Through the comparison of the results, the error

difference is small, which verifies the feasibility of this correlation method. Then, this simulation analysis method is applied to the existing project. At the same time, some data of the existing project are compared and verified by simulation and numerical value to verify the correctness of the relevant numerical value and simulation results.

4. The Influence of the Ball Screw of the Steering Gear on the Thermal Characteristics of the Box-Level Control Circuit

When the steering gear ball screw is in operation, the power loss of its driving motor, the bearing, and a certain amount of heat are generated due to friction, and the heat generated will affect the steering gear control circuit in the sealed steering gear device. Therefore, it is necessary to study the thermal response of the heat generated by the ball screw under the control circuit, and analyze the thermal response of the steering gear control circuit alone to explain the magnitude of the heat effect of the ball screw. Figure 1 shows the three-dimensional structure of the ball screw.

4.1. Calculation of Heat Source of Steering Gear Ball Screw

4.1.1. Calculate the Heat Generated by the Motor. The motor and the lead screw are connected by a coupling, so the heat of the motor is transferred to the lead screw end by heat conduction. The heat is mainly related to the output torque of the motor and the mechanical efficiency of the motor. The heating of the motor is calculated by the following formula:

$$Q = \frac{M_T \cdot n_1}{9550(1 - \eta)},\tag{1}$$

formula: Q—heat dissipated, the unit is kW; M_T—the output torque of the motor, the unit is $N \cdot m$; n_1—motor speed, the unit is r/mim; η—mechanical efficiency of the motor, generally 0.85–0.9, here it is 0.9.

4.1.2. Calculation of Bearing Friction and Heat Generation.
The friction of rolling bearings is mainly due to the heat generated by the total resistance of the components inside the bearing to the movement when the inner and outer rings of the bearing rotate. Therefore, this part of the heat is not only related to the friction torque of the bearing but also proportional to the rotational speed of the ball screw. The heat can be calculated by the following formula:

$$Q = 1.047 \times 10^{-4} n_2 M,\tag{2}$$

formula: M—frictional moment of rolling bearing; n_2—screw speed.

The friction torque M of a rolling bearing is mainly composed of two parts, the torque M_0 related to the bearing speed, type, and lubricant properties, and the friction torque M_1 related to the bearing load, $M = M_0 + M_1$. M_0 is related to the bearing speed, type, and lubricant properties, and reflects the hydrodynamic loss of the lubricant.
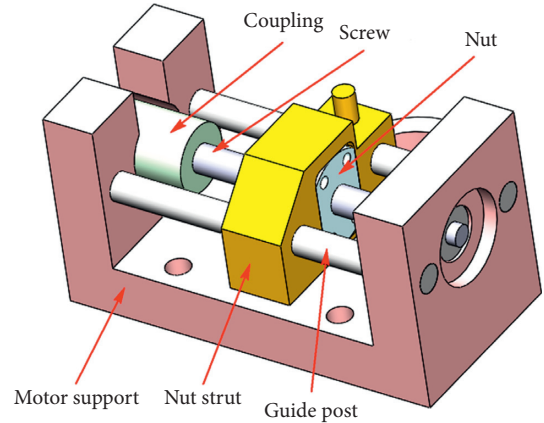
FIGURE 1: Three-dimensional structure of ball screw.

4.1.3. Ball Screw Nut Pair Calorific Value Calculation. During the operation of the ball screw, the screw rotates and drives the balls to roll along the spiral raceway. At the same time, the balls themselves are also rotating. At this time, the friction between the ball screw and the nut is expected to produce friction between the balls. The heat is the heat generated by the ball screw nut.

The calculation expression of this part of the heat is similar to the friction heat of the rolling bearing, but the friction torque of the ball screw nut pair is composed of the driving torque M_D and the resistance M_P. The driving torque M_D of the screw refers to the driving torque that the screw drives the nut to do reciprocating motion against the axial load, Which is calculated as follows:

$$M_D = \frac{F_a P_h}{2\pi\eta},\tag{3}$$

formula: F_a—axial force of the screw; P_h—lead screw; η—the working efficiency of the screw nut transmission.

The resistance torque M_P of the ball screw refers to the preload of the screw nut pair, and the calculation formula is as follows:

$$M_p = \frac{F_p P_h * \left(1 - \eta^2\right)}{2\pi},\tag{4}$$

where F_P is the axial preload of the screw.

By querying the above parameter data and calculating by formula, it can be obtained that the heating power of the driving motor is 6.713 W, the friction heating power of the nut pair is 5.867 W, and the friction heating power of the bearing is 0.3995 W. According to the analysis of the actual situation, the heat generated by the friction of the nut pair and the bearing is mostly diffused by internal heat conduction loss, and the heat exchanged by convection through the air medium is very small, so the heat of the ball screw is equivalent to the heat of the driving motor.

4.2. Analysis of the Influence of the Ball Screw of the Steering Gear on the Thermal Simulation of the Control Circuit.
The working power of the heating source of the steering gear ball screw motor is 6.7 W, and the working model of the

driving motor and the steering gear control circuit are the same. The trend is shown in Figure 2, and the motor power value is 6.7 w.

The temperature trend is obtained by iteratively calculating the power load applied to the drive motor by finite element. Figure 3 shows the maximum temperature distribution trend chart of the drive motor after the end of the working cycle. The maximum temperature increases linearly during the working time and during the cooling time. Due to the large size of the motor, it is not enough to have a significant temperature change during the cooling time of this cycle. Therefore, the maximum temperature of the driving motor progresses with the cycle, and the maximum temperature shows a stepwise increase trend, and reaches the maximum temperature at the final time 20.522°C. The temperature rise rate is only 0.00348°C/s, so it is quantitatively considered that the heat source of the ball screw has little influence on the steering gear control circuit.

5. Establishment of a Three-Dimensional Finite Element Model of the Subrack-Level Steering Gear Control Circuit Board

5.1. Finite Element Model. The physical diagram of the steering gear device is shown in Figure 4. The control circuit and the ball screw are separated on both sides and fixed with the contact surface by bolts. Figure 5 shows the steering gear box with a 2 mm heat insulation layer designed according to the physical model. It fits seamlessly with all sides of the steering gear box to form a closed body. In order to improve the calculation efficiency, the steering gear control circuit model will be further simplified, and the SOP package with less impact on the overall body and other components far away from the heat source are ignored; in the previous section, by calculating and analyzing the heat generated by the motor, the frictional heat of the ball screw, and the heat generated by the nut pair, it is known that the frictional heat of the ball screw is mainly internal solid conduction. The influence of convection heat transfer on the control circuit is small, so in order to reduce the calculation load, the ball screw is ignored in the finite element simulation analysis model in this article. The final three-dimensional finite element model of the steering gear structure is shown in Figure 6.

5.2. Material Properties. In the simulation analysis of this article, the steering gear control circuit and the steering gear ball screw are assembled in the box as the research body. At the same time, the thermal characteristics of the steering gear control circuit are explored to identify whether there is a heat insulation layer and the influence of the parameters of the heat insulation layer on the thermal characteristics of the steering gear control circuit. Material is structural steel (steel structure), an epoxy material is an insulating layer, and each of the parameter values as detailed in Tables 1 and 2.

5.3. Meshing. The whole steering gear device is mainly divided into a free tetrahedron, and the parameters of each part are customized to achieve the purpose of optimal grid
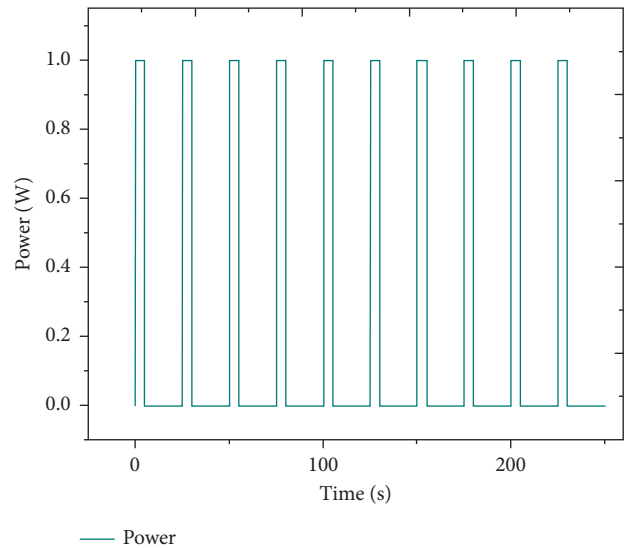
Figure 2: Power cycle load.
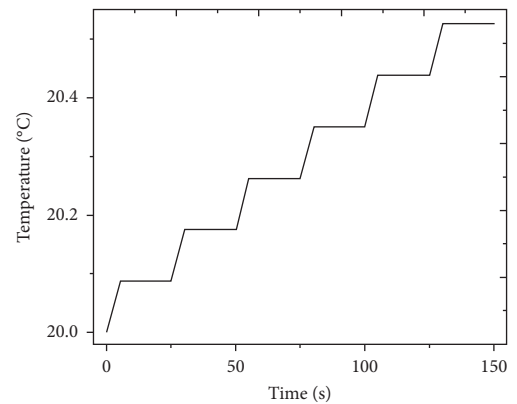
Figure 3: Trend chart of maximum temperature distribution of drive motor.
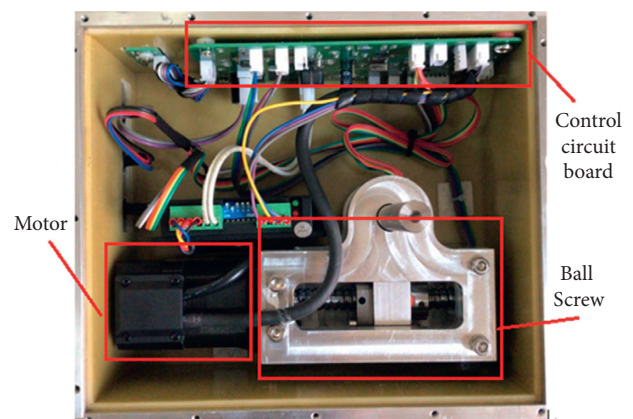
Figure 4: The physical diagram of the control circuit of the box-level steering gear.

division. Set the minimum unit parameter of the important component QFP package, the value is 0.001 mm, the minimum unit parameter of the chip power inductor is set to
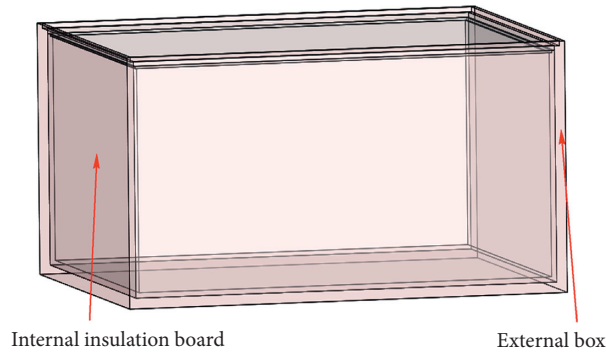
FIGURE 5: Steering gear box with heat insulation layer.
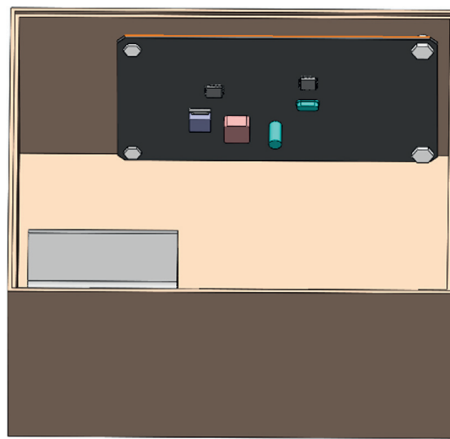
FIGURE 6: Simplified assembly drawing of the steering gear.

TABLE 1: Thermal simulation material properties.

Material	Thermal conductivity (W/(m K))	Density (kg/m³)	Specific heat capacity (J/(Kg K))
Structural steel	44.5	7859	475
Epoxy resin	0.1	980	1000

TABLE 2: Thermal-mechanical simulation material properties.

Material	Young's modulus (MPa)	Poisson's ratio	Thermal expansion coefficient [1/K]
Structural steel	2E5	0.3	$1.23E\text{-}5$
Epoxy resin	1.2E5	0.3	$6.71E\text{-}5$

0.01 mm, and the minimum unit parameters of the other PCB, steering gear, and heat insulation layer are set to the 0.1–5 mm range, the final overall meshing diagram is shown in Figure 7. Combined with the grid distribution diagram, it is found that the grid value around the important components is close to 1, and the quality is good. Therefore, it is believed that the grid division meets the accuracy requirements and can be used in finite element simulation calculations.

If the grid is further refined to make the grid quality reach about 0.85, the temperature inside the box is 70°C when it is stable, and the maximum temperature difference between the two is 1.4°C, it is considered that the difference is small. Within the acceptable range, it is quantitatively considered that the grid division with unit mass of 0.7955 has met the requirements of grid independence.

5.4. Boundary Conditions. The research and analysis process in the steering gear box is based on the working state of the rudder control circuit itself, so its boundary conditions are set as follows:

(1) Apply 1 W cyclic power to the heating source of the steering gear control circuit. Its changing trend is shown in Figure 2. In the power cycle load, the
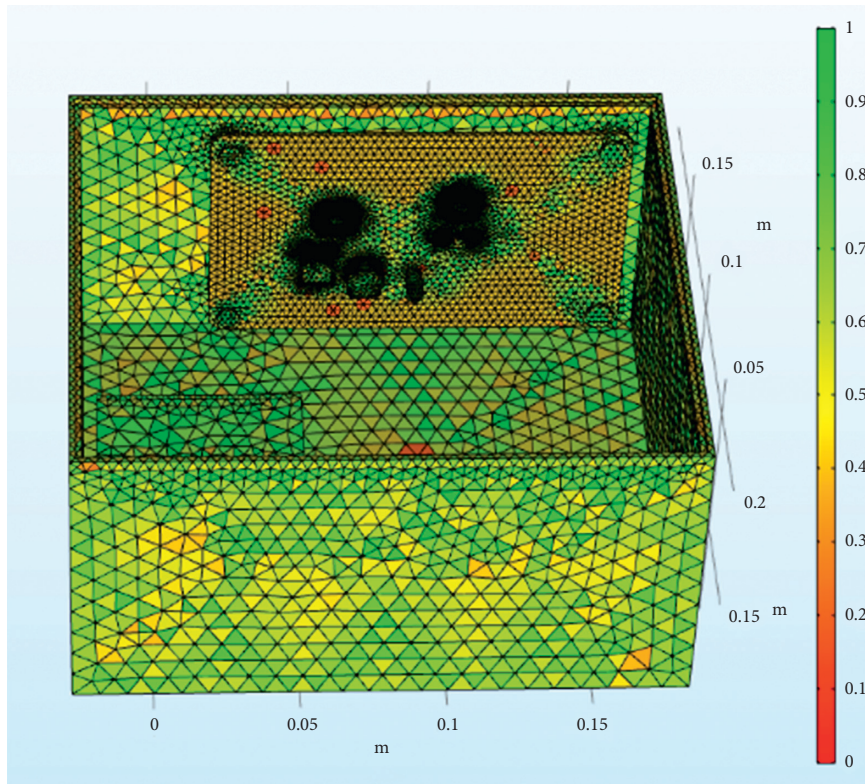
FIGURE 7: Overall finite element mesh division of the steering gear.

ambient temperature is set to 20°C, the power is increased to the rated power after 0.1 s from the beginning of the cycle, and the duration is 5 s. After that, the power is reduced from the rated power to 0 after 0.1 s, and the cooling time of 20 s is a complete cycle. A full cycle time is 25 s with a 20% duty cycle. The chip loading power is 1 W. In this section, the cycle power is loaded for a total of 150 s;

(2) The ambient temperature of the working condition is set to 20°C.

(3) Fixing constraints are imposed on the lower surface of the steering gear box and the four bolts participating in the fastening connection, and the fixing constraints are shown in Figure 8.

(4) Set the convective heat transfer coefficient on the inner and outer surfaces of the box, and the coefficient is set to 10 W/(m²·K).

6. Analysis of the Thermal Characteristics of the Thermal Insulation Layer to the Plug-In Box Level Control Circuit

6.1. Temperature Response Analysis. In the process of service, the steering gear will usually experience a harsher natural environment, which will have a certain impact on its internal structure. In severe cases, it will cause the steering gear to malfunction. Therefore, the existence of a thermal insulation layer should be considered at the beginning of the steering gear design. Reduce the probability of failure of the steering gear.
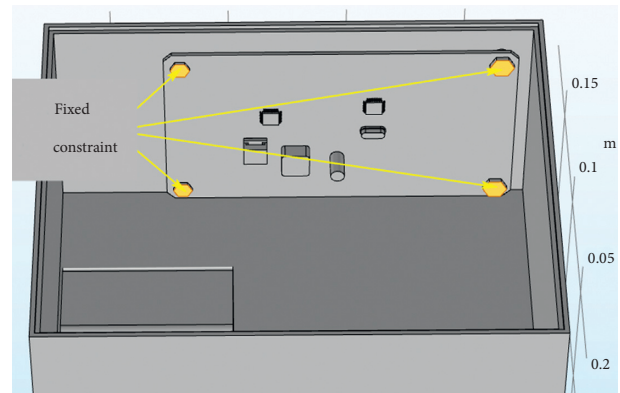
FIGURE 8: Fixed constraint setting.

In the finite element simulation calculation, the environmental temperature change load is applied to the outer surface of the steering gear. The environmental temperature was raised from the initial temperature of 20°C to 90°C over 125 s, and was maintained until 250 s, with a temperature rise rate of 0.5°C. The case surface is in contact with the insulation layer. The external heat is transferred from the case to the insulation layer through heat conduction, and finally the temperature of the inner surface of the heat insulation layer is measured. As the air flow rate inside the steering gear box is extremely low, the inner surface of the heat insulation layer is considered to be equivalent to the temperature inside the steering gear box. The measured temperature inside the steering gear box is shown in

Figure 9. It shows that the temperature inside the box reaches a stable temperature of 68.6°C, and the heat insulation efficiency reaches 24%.

The external temperature-varying load is applied to the outer surface of the steering gear, and the measured time-varying temperature inside the steering gear is equivalent to the convective heat exchange external temperature, which is used to simulate the process of air transferring heat to the control circuit through convective heat exchange. After iterative calculation, the final temperature distribution of the steering gear under the action of the heat insulation layer is shown in Figure 10. The maximum temperature appears at the contact point between the lower surface of the control circuit and the heat insulation layer. The maximum temperature value is the ambient temperature of 90°C, and the minimum temperature value of 23.8°C appears on the drive motor. The drive motor is made of copper and has a large volume. At the final moment, there has not been a sufficiently high temperature rise, and the temperature value is low.

The steering gear control circuit is the core control component of the overall steering gear. Figure 11 is the cloud map of the temperature distribution of the PCB board at the final moment under the action of the thermal insulation layer. The main heat sources of the PCB include three parts: including the heat transferred through the convective heat transfer of the air medium. The bolt is in direct contact with the steering gear box and the control circuit through the heat transferred by heat conduction, and the heat generated by the intermittent operation of the heating source on the control circuit is transferred to the PCB through the pin solder foot heat. Therefore, it can be found in the figure that the maximum temperature of the control circuit appears at the bottom right corner where the lower surface is in contact with the heat insulation layer and reaches 65.7°C, and the minimum value appears on the upper surface of the control circuit, and its value is 50°C. The lower surface temperature of the control circuit is higher than the others. Components and the four bolt holes are relatively high due to the presence of heat conduction. The heat generated by the chip and the patch power inductor is relatively small compared to the outside temperature, so there will be no obvious temperature gradient near it.

Figure 12 shows a QFP package pin and a temperature change of the fillet. The curve is derived from the QFP package with increasing time. The temperature is gradually increased. At the time, the final temperature reached 51°C. At the same time, small periodic bulges will appear during the overall temperature rise of QFP package. This is because during this period of time the chip generates a certain amount of heat due to its own power, which causes a sudden temperature rise in the corresponding period of time, resulting in bumps. In addition, the heat generated by the power of the chip is transferred from the chip to the pins and then to the solder feet. The transfer process causes a certain temperature difference between the solder feet and the pins. Therefore, the temperature of the pins is slightly greater than the temperature of the solder feet during the relevant cycle period. Since the initial temperature value of QFP is low, the temperature rise rate is lower in the early stage of the environmental temperature effect, and the curve climbing trend is slower. In the later stage of the effect, the temperature rise rate is faster and the curve climbing trend is steeper.

6.2. Structural Response Analysis. Import the above temperature results into the structure module to calculate the structure results, as shown in Figure 13, which is a cloud diagram of the stress distribution of the steering gear at the final moment under the action of the heat insulation layer. Because the ambient temperature of the steering gear box and the steering gear itself has a larger temperature gradient, the wall of the steering gear produces a larger stress value, and the maximum stress value is at the edge position where the cover plate contacts the case body, which is a stress concentration phenomenon.

Since the overall stress gradient distribution of the steering gear is quite different, the PCB and the four bolts are separately analyzed for stress. As shown in Figure 14 is the stress distribution cloud diagram of the PCB board, the maximum stress appears on the surface of the bolt hole in the upper left corner, and the value is 75.3 MPa; Figure 15 is a cloud diagram of the stress distribution of four bolts. The maximum stress appears on the surface of the bolt at the lower right corner, and its value is 316 MPa. Query data learned by, the PCB yield strength of 350 MPa, yield limit of the steel of 454 MPa, the PCB and the bolt stress values are less than the corresponding yield limit.

Figure 16 shows the stress distribution trend diagram of the first group of pins and solder feet of the QFP package. The change trend is similar to the temperature change trend, and the stress value gradually increases with time. Among them, the copper pin stress change trend has obvious bump changes with the change of the chip power cycle, and the tin-lead soldering pin changes smoothly with the cycle. At the final moment, the stress value of the copper pin reached 90.7 MPa, and the stress value of the tin-lead solder pin reached 41 MPa.

Through the analysis of the stress value of PCB, four bolts and QFP pin solder foot, combined with the corresponding yield limit of each material and their respective common failure modes, it is conservatively considered that the common failure location of the steering gear control circuit is QFP under severe environmental conditions—the connection point between the pin and the solder foot.

7. Analysis of Thermal Characteristics of Sub-Box-Level Control Circuit with or without Heat Insulation Layer

In the previous section, the thermal response results of the steering gear box and the internal control circuit under the action of the heat insulation layer were elaborated. This section quantitatively analyzes the influence of thermal insulation by comparing the thermal characteristic response under the action of thermal insulation. The lower surface of the internal control circuit of the thermal insulation model is
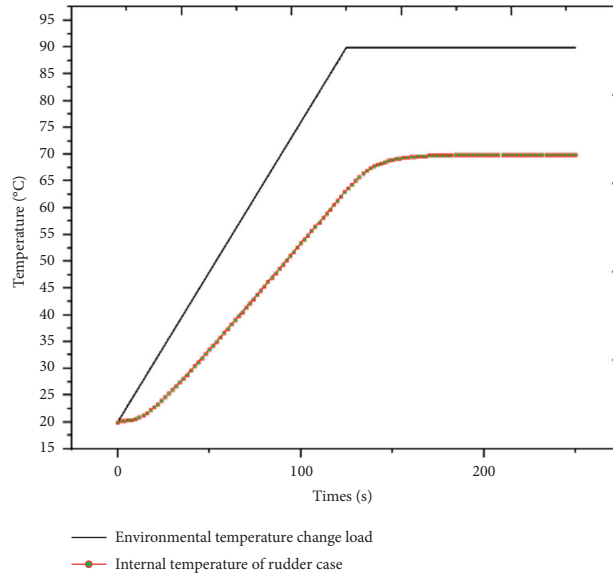
FIGURE 9: Internal and external temperature curve of steering gear.
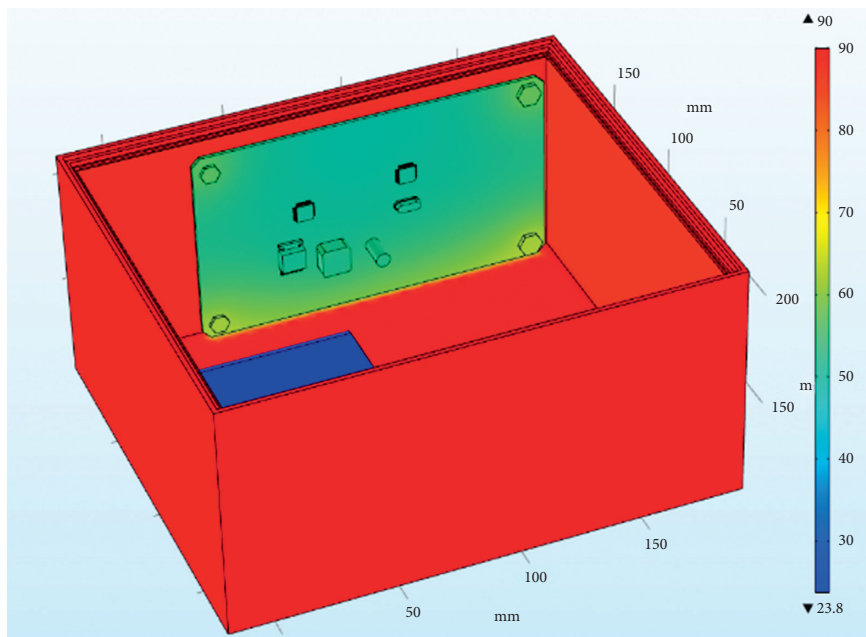
FIGURE 10: Cloud diagram of the temperature distribution of the steering gear at the final moment under the action of the heat insulation layer.

in contact with the thermal insulation, while the model without the thermal insulation is in contact with the surface of the box. In the finite element analysis and calculation, Figure 17 shows the internal temperature curve of the cabinet with or without the heat insulation layer. From the figure, it can be seen that the internal temperature value without the heat insulation layer is close to the external ambient temperature. When it reaches 89.8°C, the internal temperature value under the action of the heat insulation layer is 68.6°C. Only from the internal temperature performance analysis of the box, the thermal insulation

efficiency of the thermal insulation layer is 24%, and the thermal insulation effect is significant.

The external temperature change load is applied to the outer surface of the steering gear, and the measured time-varying temperature inside the steering gear is equivalent to the external temperature of convective heat transfer. Figure 18 shows the overall temperature distribution cloud diagram with or without heat insulation. In the figure, it can be seen that the overall temperature of the internal control circuit without the effect of the heat insulation layer is significantly higher than that of the internal control circuit
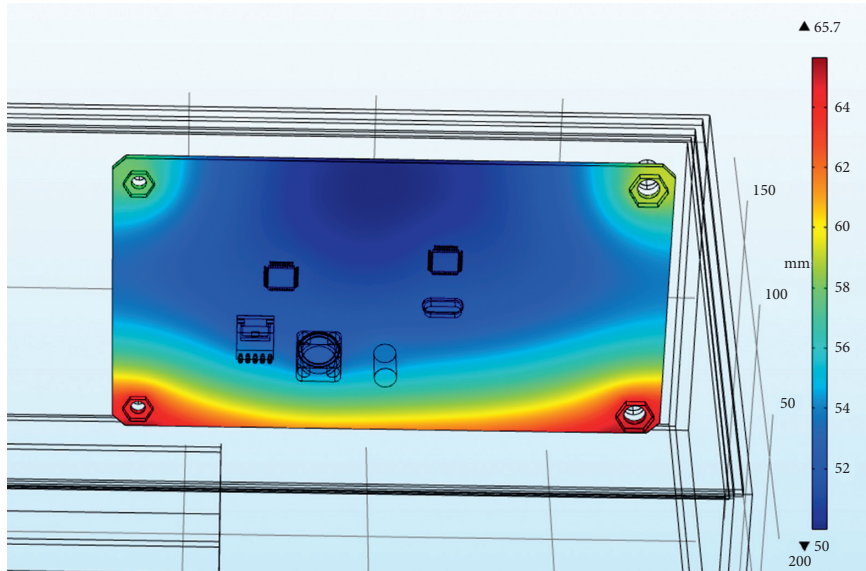
FIGURE 11: The temperature distribution cloud diagram of the PCB board at the final moment under the action of the heat insulation layer.
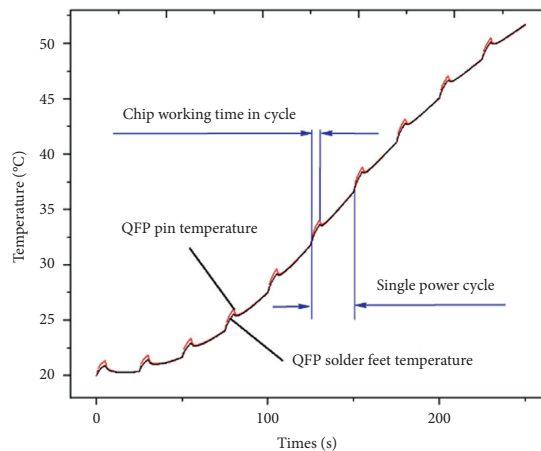
FIGURE 12: QFP package body pin and solder foot temperature change trend.

under the effect of the heat insulation layer. The two models have similar temperature distribution trends on the control circuit. The temperature of the four bolts is higher than that of other parts of the control circuit board, and the temperature of the middle position is lower. As shown in Figure 19, the temperature distribution of PCB with or without insulation layer is completely consistent. The maximum temperature of PCB without insulation layer is 90°C, and the minimum temperature is 68°C. The average temperature of the PCB board is 79°C; the highest temperature of the PCB board under the action of the heat insulation layer is 65.7°C, the minimum temperature value is 50 °C, and the average temperature of the PCB board is 57.4°C. The overall temperature of the two models is quite different.

Figure 20 shows the maximum temperature curve of the PCB board with or without the heat insulation layer. In the model without the heat insulation layer, since the lower surface of the control circuit and the bolt surface are directly

in contact with the steering gear box, the structural steel material has good thermal conductivity; therefore, the temperature rise rate of the PCB board is very fast, reaching the maximum value of 89.2 °C in a short time, and the temperature change curve is steep; in the insulation layer model, due to the existence of the insulation layer, the temperature that reaches the inside within the same time is much less than where there is no insulation layer. In the thermal layer model, the maximum temperature at the final moment is 65.7 °C, and from the figure, it can be concluded that the temperature growth trend of the PCB board is relatively flat.

Figure 21 shows the temperature curves of the three bolt faces with or without the heat insulation layer. The analysis shows that the temperature values of the left and right bolts on the lower surface are similar. Here, we only study the temperature of the three bolt faces. The analysis shows that there is no heat insulation. The bolt temperature value and the temperature rise rate in the layer model are greater than those in the insulation layer model. A comparative analysis
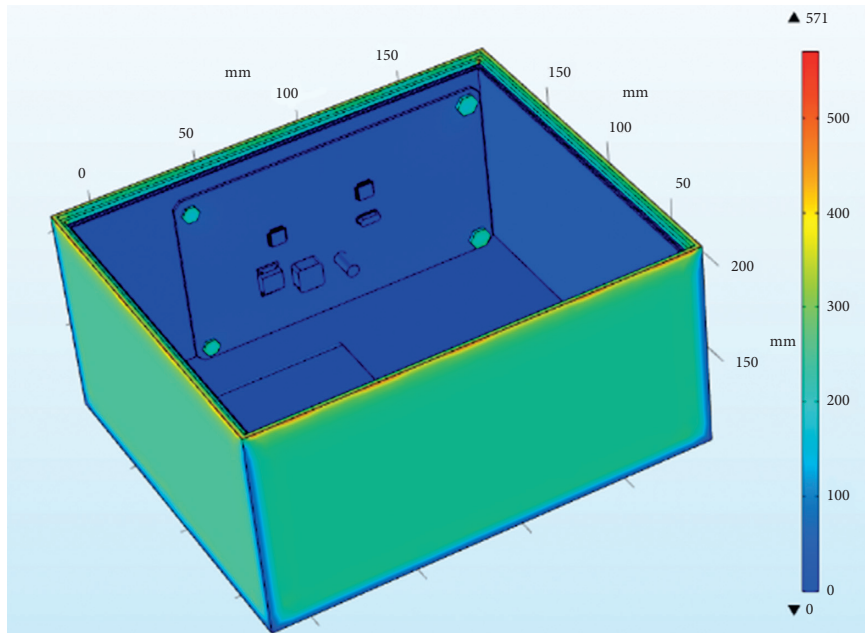
FIGURE 13: The stress distribution cloud diagram of the steering gear at the final moment under the action of the heat insulation layer.
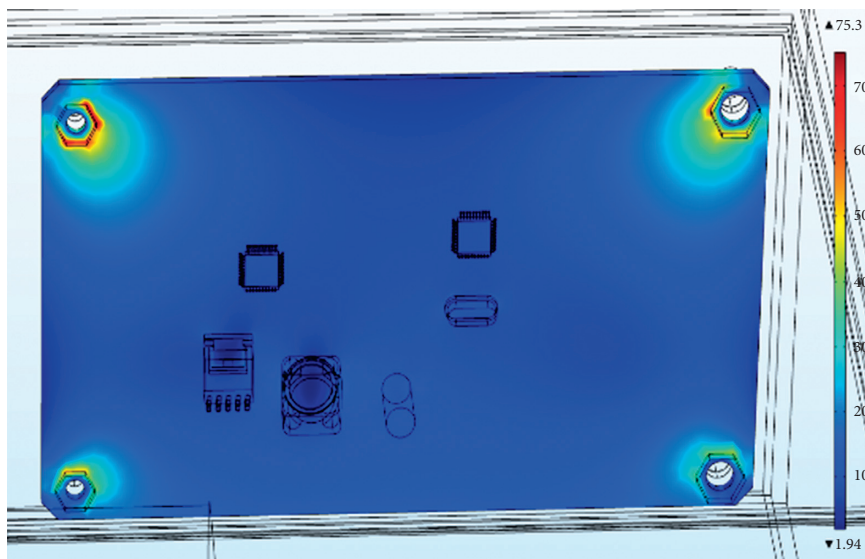
FIGURE 14: The stress distribution cloud diagram of the PCB board.

of the maximum temperature changes of PCB boards and bolts in the model with or without the insulation layer is carried out. The temperature and rate of temperature rise of the insulation layer model are lower than those of the model without insulation layer, and the values are quite different.

Figure 22 shows the temperature curve of QFP pin solder fillet with and without thermal insulation layer. The temperature value of QFP pin solder pin under the thermal insulation layer model and the QFP pin solder pin under the model without thermal insulation layer are almost the same. At the same time, it can be found that the temperature values at the same location in the two models are quite different.

Figure 23 shows the overall stress cloud diagram with or without the thermal insulation layer. It can be concluded that stress concentration occurs in the overall stress distribution with or without the thermal insulation layer. Among them, the overall stress without the thermal insulation layer is greater than that with the thermal insulation layer. With the overall stress under the action, the maximum stress on the control circuit board inside the box still appears on the bolt surface.

The QFP pin soldering foot of the dangerous component on the control circuit is selected as the research object. Figure 24 shows the stress of the QFP pin soldering foot with or without the heat insulation layer. The final stress values of the pins without the heat insulation layer are 153 MPa and 90 MPa, respectively; the final stress values of the pins under the action of the heat insulation layer are 61 MPa and 41 MPa, respectively. The solder foot is made of tin-lead
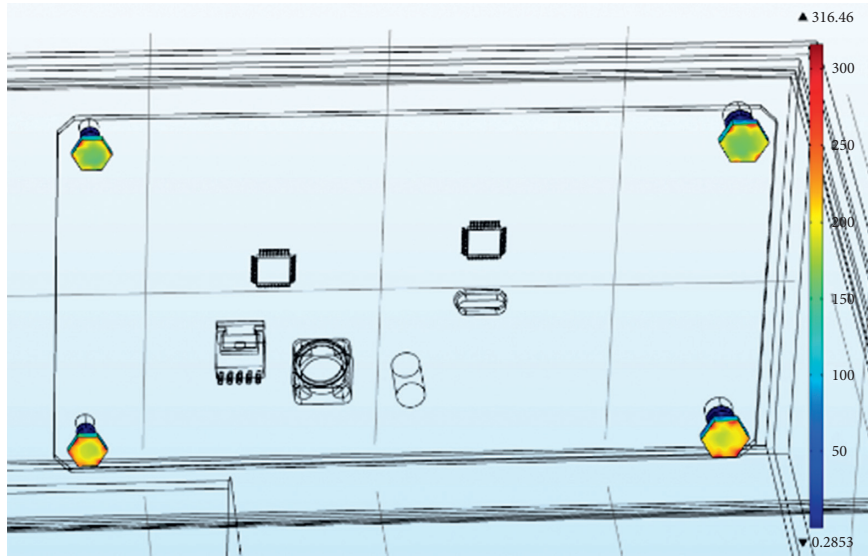
FIGURE 15: Stress distribution cloud diagram of four bolts.
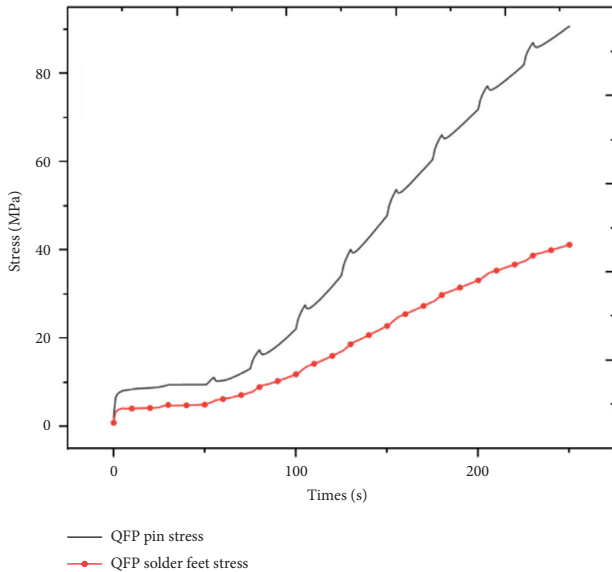
FIGURE 16: QFP package body lead and solder foot stress trend.

FIGURE 17: The internal temperature curve of the box with or without the heat insulation layer.

material, and the required stress value is 45 MPa. Therefore, the solder foot failure phenomenon may occur without the action of the heat insulation layer, which will cause the steering gear to fall into failure mode. By comparing the temperature and stress of the PCB board, bolts, and QFP pin solder feet under the action of the heat insulation layer, the values of the two are quite different, and the effect of the heat insulation layer is significant. Therefore, it is considered that heat insulation should be considered at the beginning of the design. Layers are necessary and indispensable.

Through comparative analysis of the temperature and stress response results of PCB board, bolts, and QFP pin solder feet under the action of the heat insulation layer, the response value under the action of the heat insulation layer is much smaller than the corresponding value under the action where there is no heat insulation layer In response to the

numerical value, the effect of the insulation layer is significant, so it is considered necessary to consider the insulation layer at the beginning of the design.

8. Comparative Analysis of Thermal Characteristics of Different Parameters of Insulation Layer

8.1. Comparative Analysis of Thermal Properties of Different Thermal Conductivities of Thermal Insulation Layers. The thermal insulation layer of different materials has different thermal conductivity, and the heat transfer from the external environment to the inside through the heat conduction method is different, which makes the temperature gradient generated by
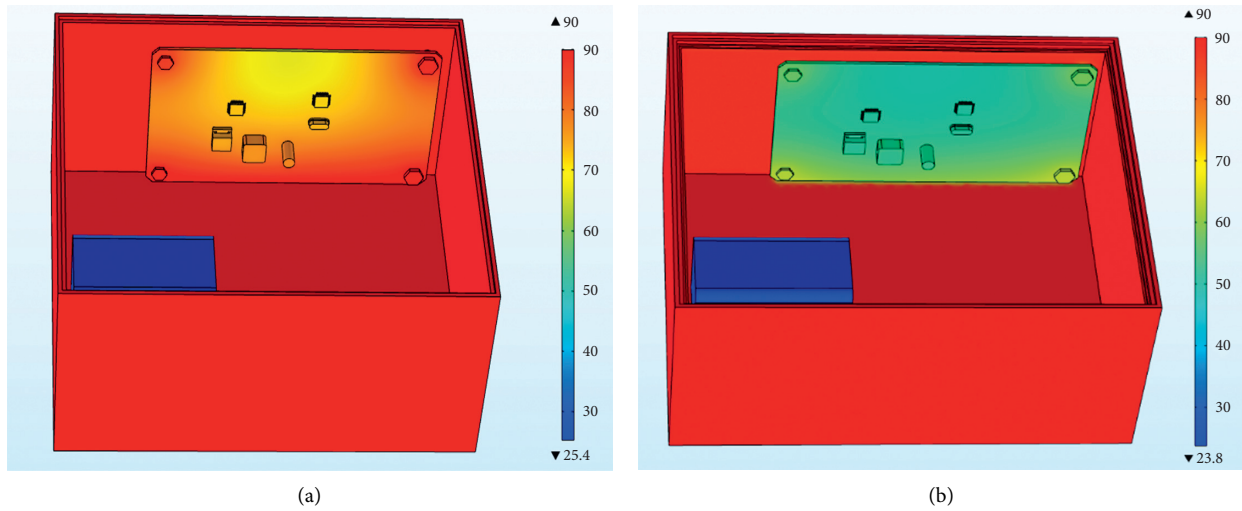
FIGURE 18: Cloud diagram of overall temperature distribution with or without heat insulation layer. (a) Cloud diagram of overall temperature distribution without heat insulation layer. (b) Cloud map of the overall temperature distribution under the action of the thermal insulation layer.
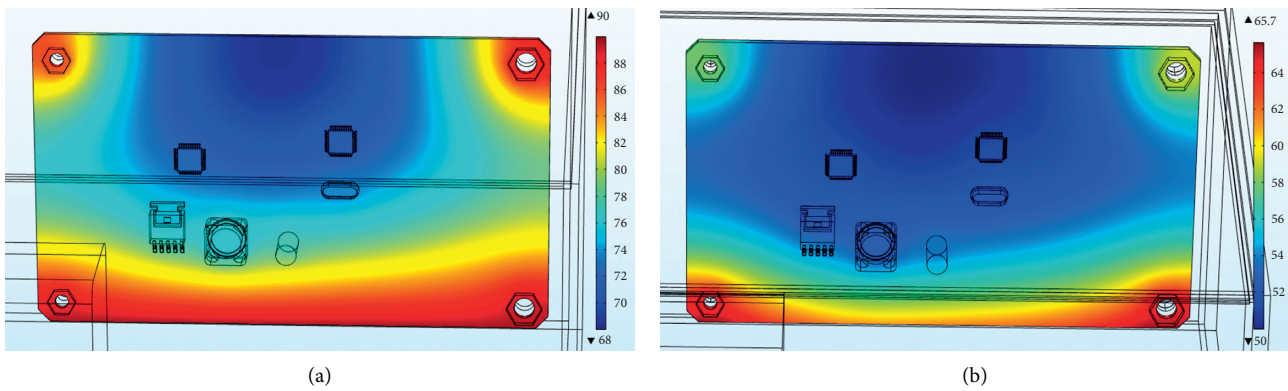
FIGURE 19: PCB board temperature distribution with or without heat insulation layer. (a) PCB board temperature distribution without heat insulation layer. (b) PCB board temperature distribution under the action of the heat insulation layer.
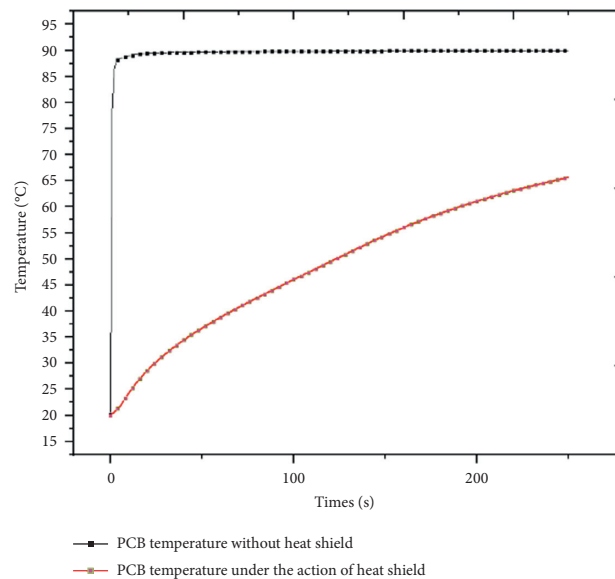
FIGURE 20: PCB board temperature with or without the heat insulation layer.
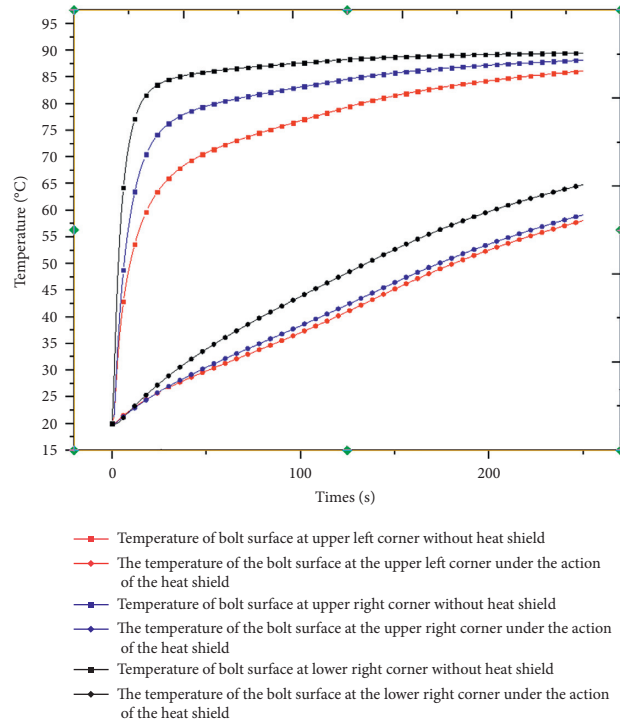
Figure 21: Bolt surface temperature with or without the heat insulation layer.
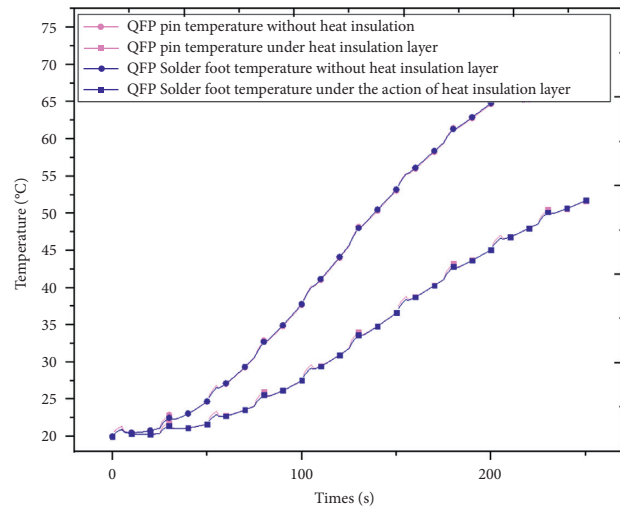
Figure 22: QFP pin solder foot temperature with or without heat insulation layer.

the internal heat acting on the control circuit different. This section aims to study the thermal response of different materials and selects the one with better comprehensive performance of the selected materials by comparative analysis.

This section selects foamed polyurethane (EPU) materials with a thermal conductivity of 0.047 and asbestos materials with a thermal conductivity of 0.152, and comprehensively compares and analyzes the epoxy resin materials with a thermal conductivity of 0.1 studied in the previous section, from the aspects of thermal insulation effect and economic cost. Consider choosing a more suitable insulation material.

Figure 25 shows the internal surface temperature curve measured under the action of different insulation layers when the external environmental load is applied to the box body. The internal temperature reached is also different due to the different thermal conductivity of each material. Under the action of EPU, epoxy resin, and asbestos materials, the internal temperature reaches 58.8, 69.9, and 75.1, respectively. Due to the certain hysteresis of heat conduction, the time node when the internal temperature reaches stability under the action of the three thermal insulation materials is greater than the time node when the ambient temperature reaches stability.

(a)   (b)
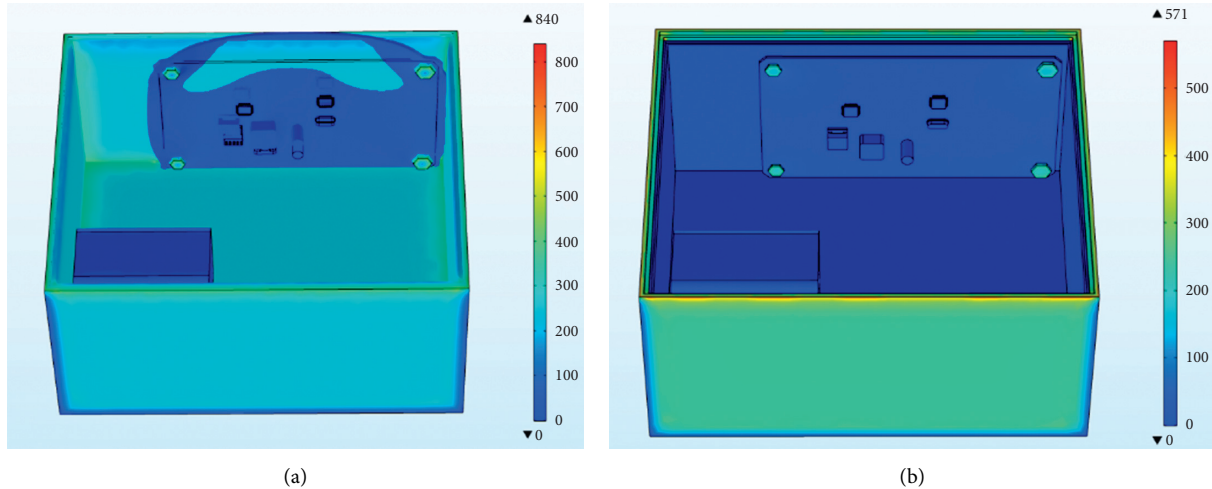
FIGURE 23: Overall, stress cloud diagram with or without heat insulation layer. (a) Overall stress cloud diagram without heat insulation layer. (b) The overall stress cloud diagram under the action of the insulation layer.
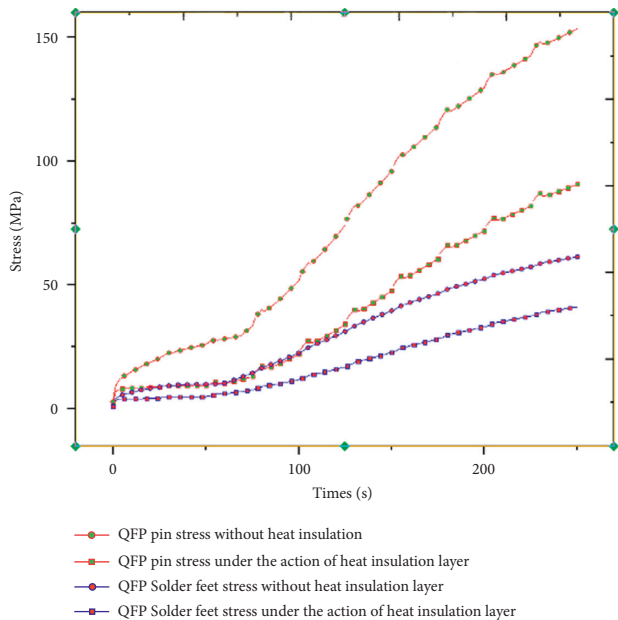
FIGURE 24: QFP pin solder foot stress with or without heat insulation layer.
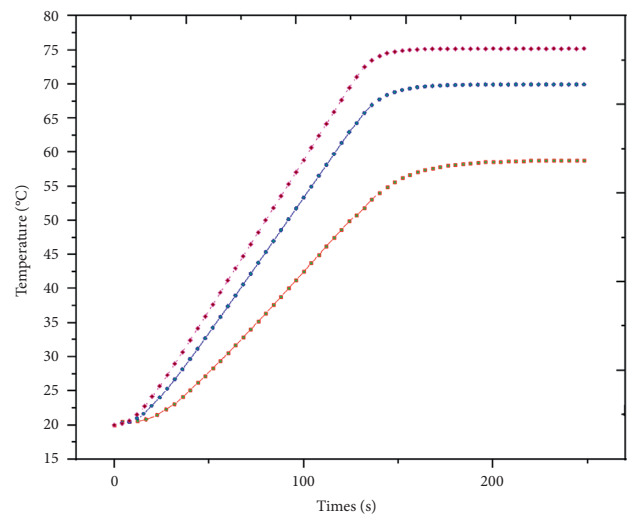
FIGURE 25: The internal temperature of the box under different materials.

Figures 26(a)–26(c) are the overall temperature cloud diagrams under EPU, epoxy resin, and asbestos materials, respectively. It can be found from the figure that the overall temperature of the internal control circuit under the action of the EPU heat insulation layer material is the lowest, and in the asbestos the overall temperature of the internal control circuit under the action of the heat insulation layer material is the highest. With the increase of the thermal conductivity, the temperature of the internal control circuit increases. The temperature trend under the action of the three materials is the same, and the temperature value of the four bolts is higher than that of other positions of the PCB.

The control circuit is located in the steering gear and exchanges heat with internal heat through convection heat exchange. The temperature rise trend of different components on the control circuit is analyzed separately, and PCB and QFP solder feet are selected to study the temperature rise trend. Figure 27 shows the temperature rise curve of the PCB board under the action of three different heat insulation materials. The corresponding materials are EPU, epoxy resin, and asbestos. The maximum temperature of the PCB board at the final moment is 60.9°C, 65.6°C, and 72.7°C. The maximum temperature increase was 7.8% and 10.8% in sequence. Figure 28 shows the temperature rise curve of QFP soldering feet under the action of different insulation materials, corresponding to EPU, epoxy resin, and asbestos. The maximum temperature of QFP soldering feet at the final moment is 44.6°C, 51.7°C, and 56.2°C. The maximum temperature increase was 15.9% and 8.7% in sequence. The PCB board and QFP

(a)

(b)

(c)
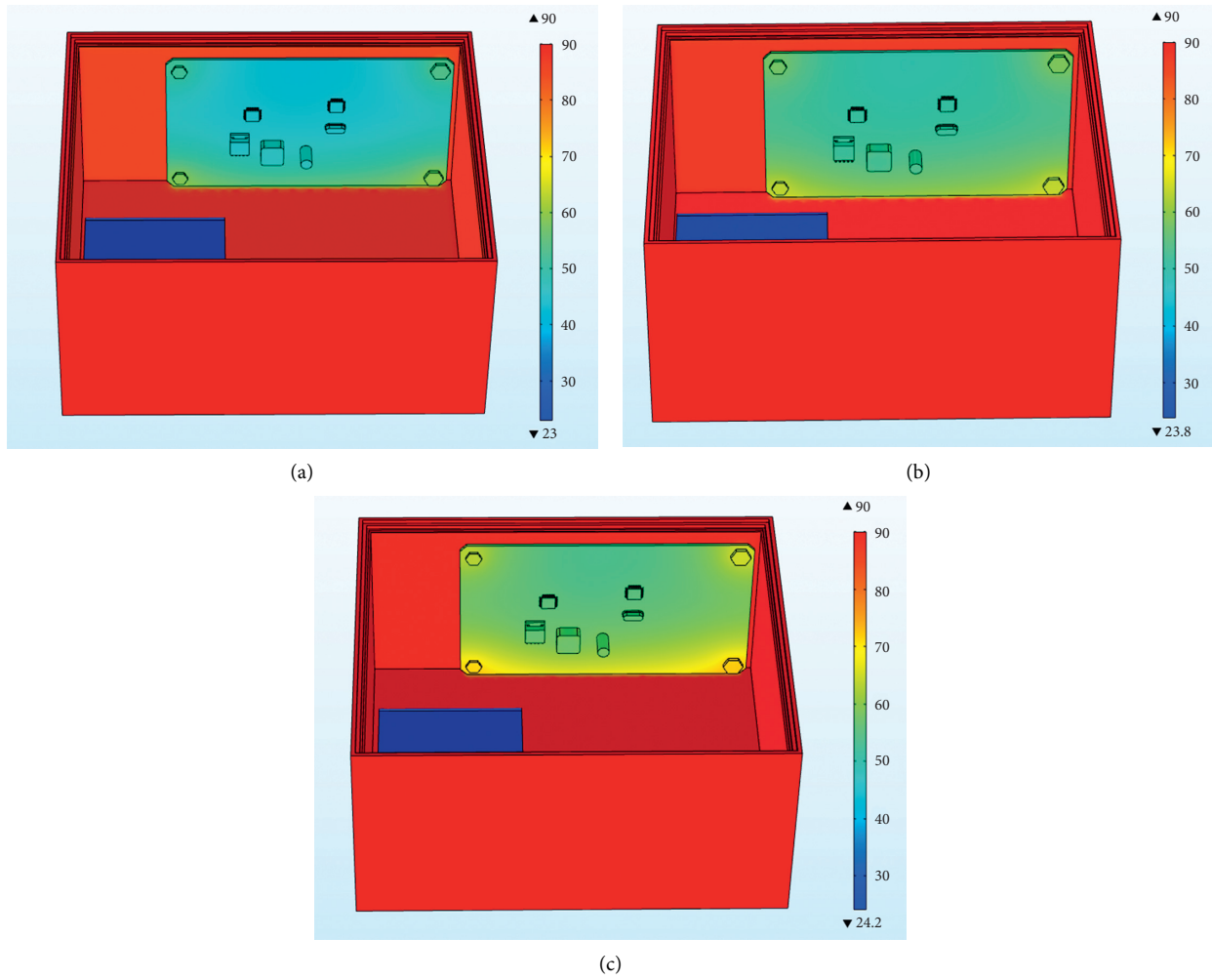
FIGURE 26: Overall temperature cloud diagram under different materials. (a) Overall temperature cloud diagram under EPU material. (b) Overall temperature cloud diagram under epoxy resin material. (c) Overall temperature cloud diagram under asbestos material.
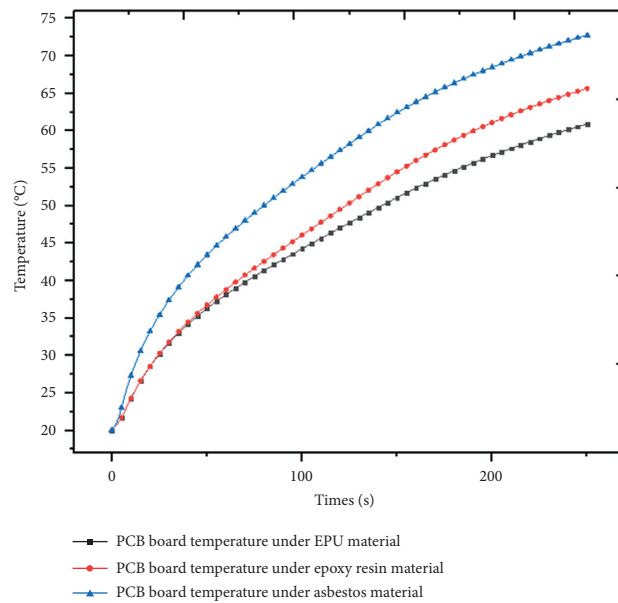
FIGURE 27: PCB board temperature under the action of different insulation materials.
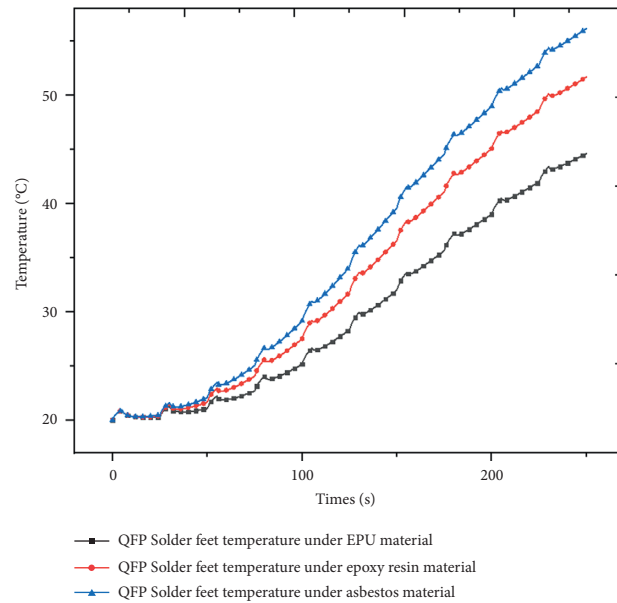
Figure 28: Welding foot temperature under the action of different insulation materials.

comparative analysis of the temperature rise curve of the solder foot shows that the temperature rise curve rate of the PCB board is first large and then small, and the temperature rise rate of the QFP solder foot is first small and then large. The reason is that the temperature of the PCB increases through the heat conduction method. The main heat source of the QFP is Convection heat exchange with the internal air causing the temperature to rise.

Since the overall stress cloud chart trends under the action of different insulation materials are completely consistent, the QFP solder foot with obvious stress law is selected as the research object for stress analysis. Figure 29 shows the stress curve of QFP welding foot under the action of different insulation materials. For EPU, epoxy resin, and asbestos, the maximum stress value of QFP welding foot at the final moment is 32.6 MPa, 41.2 MPa, 46.6 MPa, respectively, and the maximum stress increase was 26.9% and 13.1% in sequence. The stress change curve of the welding foot under the action of different insulation layer materials is similar to the temperature change curve, but the stress change is more significant than the temperature change.

Through the temperature and stress comparison analysis of different components of the control circuit under the action of different insulation materials, it is found that the thermal conductivity of the EPU insulation material is the smallest compared with the other two materials, and the insulation effect is the best, but the EPU material is more expensive in the market. Under the action of asbestos as the insulating layer material, the stress value of the QFP soldering leg reaches 46.6MPa; at the final time, the soldering leg is made of tin-lead, and its allowable stress value is about 45MPa. The internal control circuit may fail under the action of the insulating layer material asbestos. Therefore, considering the heat insulation effect and economic cost, the epoxy resin material is selected as the heat insulation material for this product.

8.2. Comparative Analysis of Thermal Characteristics of Different Thickness Insulation Layers.

For certain insulation layer materials, different insulation layer thicknesses have different thermal resistances, so that the internal temperature under the action of insulation layers with different thicknesses is not nearly the same. This section studies the insulation layer under the condition that the insulation layer material has been determined for the influence of thickness; the optimal thickness of the thermal insulation layer is also selected from various factors such as thermal insulation effect and cost. This section selects the epoxy resin insulation material determined in the previous section to study the internal control circuit under the action of three different insulation layers of thickness 2 mm, 3 mm and 4 mm, and compares and analyzes the corresponding responses of their important components and comprehensively considers various aspects. Determine the appropriate insulation thickness.

Figure 30 shows the internal surface temperature curve measured under the action of different insulation layers when the external environmental load is applied to the outside of the box. As the thickness of the thermal insulation layer increases, the thermal resistance increases, and the temperature of the inner wall of the box decreases. The temperature of the inner wall of the thermal insulation layer with a thickness of 2 mm is $69.9°C$ at the final moment, and the temperature of the inner wall of the thermal insulation layer with a thickness of 3 mm is $63.6°C$ at the final moment. The inner wall temperature of the 4mm thickness insulation layer is $58.0°C$ at the final moment, and the inner wall temperature of the box decreases by about 9% for every 1mm increase in the thickness of the insulation layer.

Figures 31(a)–31(c) are the overall temperature cloud diagrams under the action of 2 mm, 2 mm, and 4 mm insulation layers, respectively. As the thickness of the insulation layer increases, the temperature of the internal control circuit decreases. Increasing the thickness of the thermal insulation layer helps to reduce temperature of the control circuit and reduce the probability of failure.
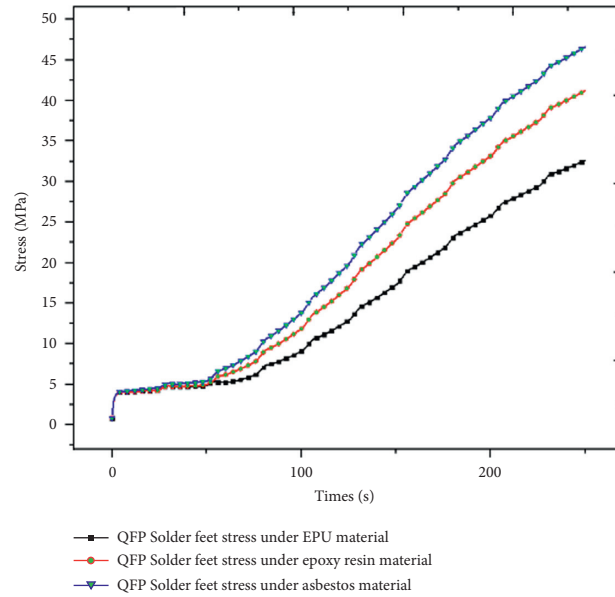
FIGURE 29: QFP welding foot stress under different heat insulation materials.
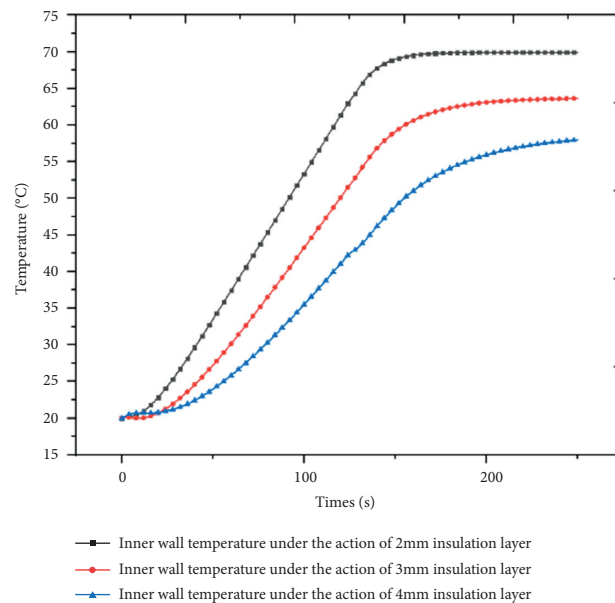
FIGURE 30: Internal temperature under the action of different thicknesses of the heat insulation layer.

Figure 32 shows the temperature rise curve of the PCB board under the action of three different thicknesses of the heat insulation layer. Under the action of 2 mm heat insulation layer, 3 mm heat insulation layer, and 4 mm heat insulation layer, the maximum temperature value of the PCB board at the final moment is 65.6°C, 61.5°C, and 58.3°C, with the gradual increase of the thickness of the thermal insulation layer, and the maximum temperature drop ratios are 10.6% and 10.6% in turn. Figure 33 shows the temperature rise curve of QFP welding feet under the action of different thicknesses of heat insulation layers. Under the actions of 2 mm heat insulation layer, 3 mm heat insulation layer, and 4 mm heat insulation layer, the maximum temperature value of QFP solder feet at the final moment is 51.7°C, 46.2°C, 41.3°C, with the gradual

increase of the thickness of the thermal insulation layer, and the maximum temperature drop ratios are 6.3% and 5.2% in turn.

Figure 34 shows the stress curve of QFP welding foot under the action of different thicknesses of the heat insulation layer. Under the action of 2 mm heat insulation layer, 3 mm heat insulation layer, and 4 mm heat insulation layer, the maximum stress value of QFP solder foot at the final moment is 41.2 MPa, at 33.9 MPa, and 26.8 MPa, with the gradual increase of the thickness of the thermal insulation layer, and the maximum stress drop ratios are 17.7% and 20.9% in turn. The influence of stress changes under the action of different thickness insulation layers is greater.

Through the comparison of the overall temperature under the action of different thicknesses of the heat

(a)

(b)

(c)
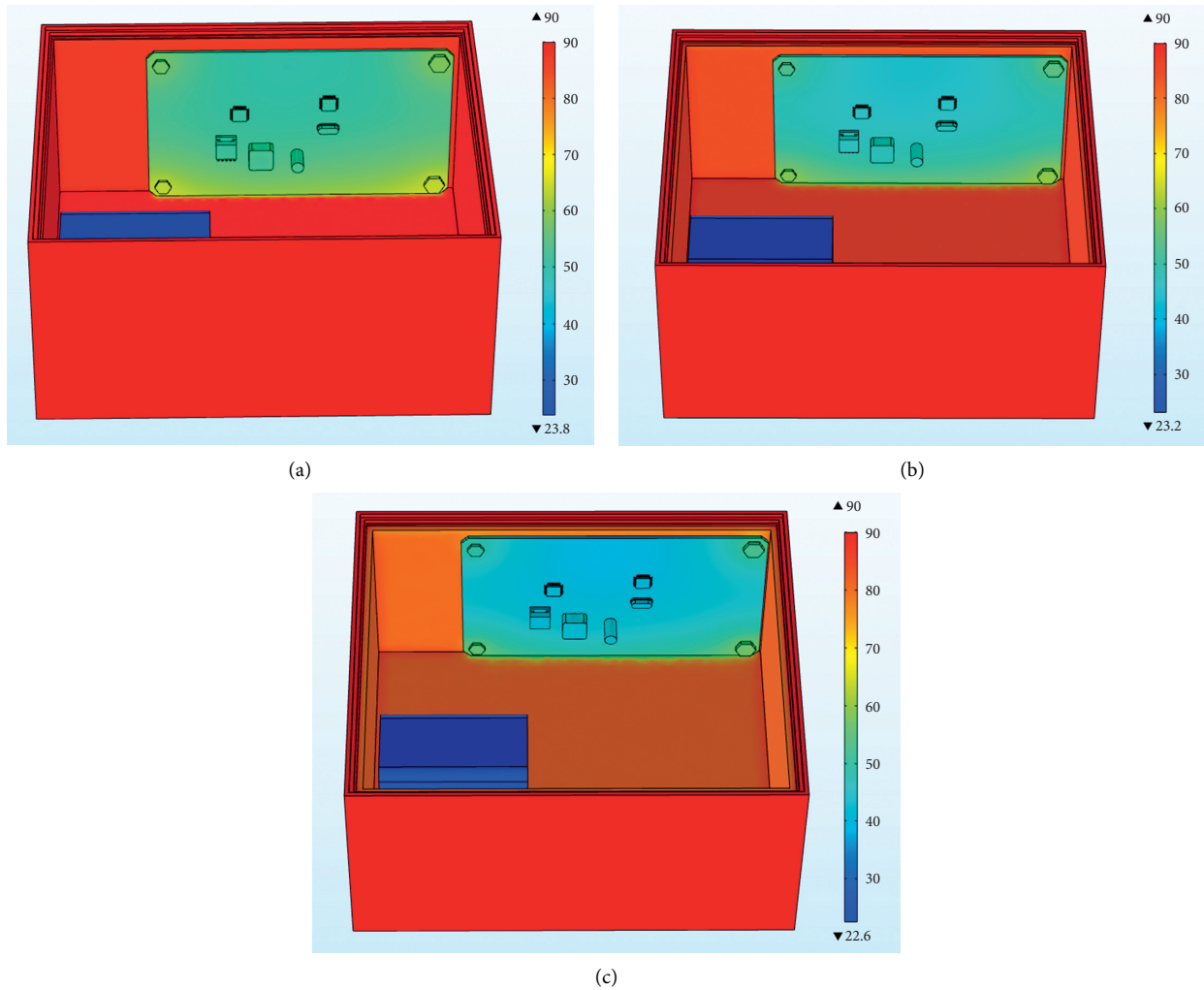
FIGURE 31: Temperature cloud diagram under the action of insulation layers of different thicknesses. (a) Temperature cloud diagram under the action of 2 mm insulation layer. (b) Temperature cloud diagram under the action of 3 mm insulation layer. (c) Temperature cloud diagram under the action of 4 mm insulation layer.
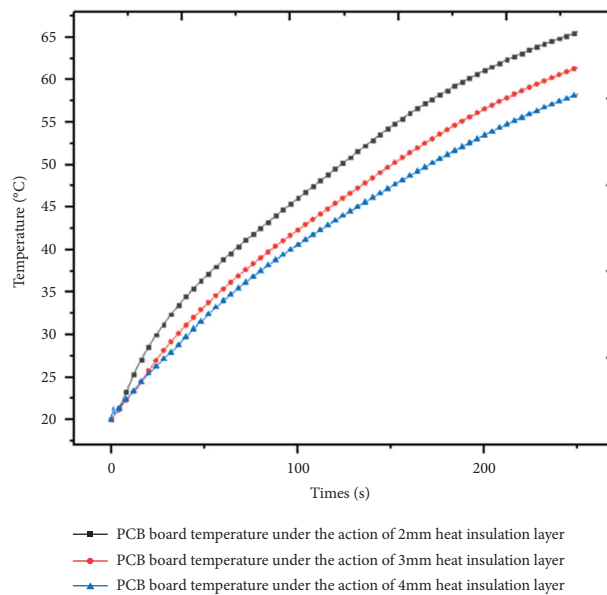
FIGURE 32: PCB board temperature under the action of different thicknesses of the heat insulation layers.
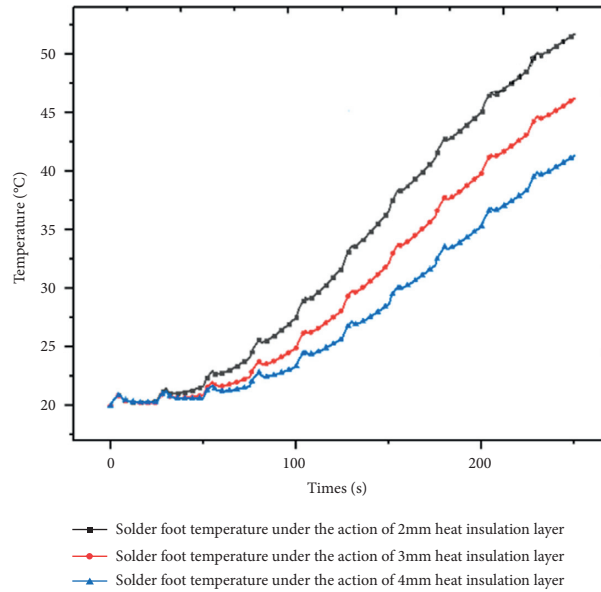
FIGURE 33: QFP solder foot temperature under the different thicknesses of the heat insulation layers.
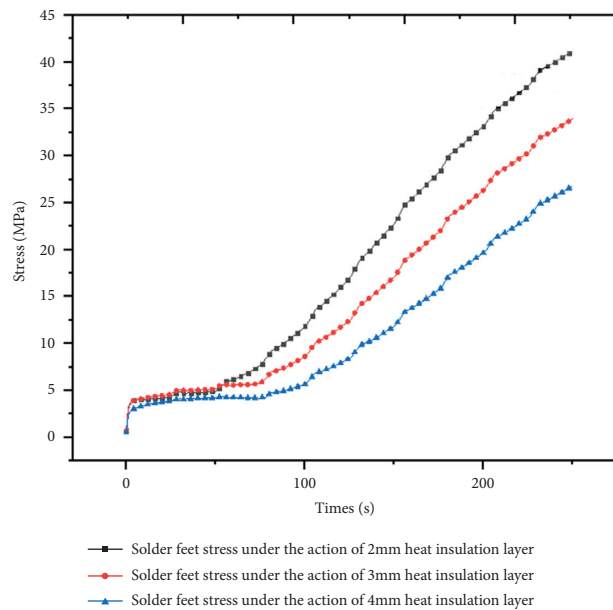
FIGURE 34: QFP welding foot stress under different thicknesses of the heat insulation layer.

insulation layer and the temperature of different components of the control circuit, it is known that as the thickness of the heat insulation layer increases, the internal temperature and heat act on the control circuit through convective heat transfer. The resulting temperature will decrease accordingly. By inquiring the electronic manual, the working permission temperature of related components on the steering gear was obtained. When the thickness of the heat insulation layer is determined to be 2 mm, the internal ambient temperature of the box has already met the working permission temperature.

Therefore, from the cost point of view, the thickness of the insulation layer of 2 mm thickness has met the working requirements; at the same time, it is known from the above analysis that the stress is more sensitive under the action of the insulation layer of different thickness, and the stress changes greatly. For high requirements, a 4-mm thick thermal insulation layer or even a larger thermal insulation layer can be used.

9. Conclusion

This paper establishes an overall three-dimensional finite element model that integrates the steering gear control circuit and the steering gear ball screw in the box, studies the influence of the internal ball screw heating source on the control circuit, and analyzes and calculates that the heat source of the ball screw is relatively high. The impact on the control circuit is small.

(1) In order to simulate the effect of the heat insulation layer of the steering gear box in the harsh natural environment during the service process, the case body is placed at a gradual temperature and finally kept at a constant high temperature environment, and a comparative analysis is made from the three perspectives of temperature and stress. Based on the finite element response of two models with and without the insulation layer, it is concluded that the effect of the insulation layer is significant, and the insulation layer is necessary in the process of design and use;

(2) Through the analysis of the thermal response of the steering gear under the conditions of different materials and different thicknesses of the heat insulation layer, the design of the heat insulation layer is further optimized and recommended.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

References

[1] X. P. Zhang, L. M. Yin, and C. B. Yu, "Research and application progress of lead-free solders for electronic and photonic packaging," *Journal of Materials Research*, vol. 22, no. 1, pp. 1–9, 2008.

[2] R. Ye, G. X. Guan, and H. S. Guan, "Analysis of the influence of heat transfer path of high heat flux density chip," *Modern manufacturing*, no. 24, pp. 110–113, 2017.

[3] Z. W. Zhong, K. W. Wong, and X. Q. Shi, "Interfacial behavior of a flip-chip structure under thermal testing," *IEEE Transactions on Electronics Packaging Manufacturing*, vol. 27, no. 1, pp. 43–48, 2004.

[4] S. X. Su, X. G. Du, and Y. B. Qiao, "Research on thermomechanical stress characteristics of integrated circuit chips," *Electronics World*, vol. 21, no. 11, pp. 26-27, 2014.

[5] W. Liu, L. Y. Yu, and S. Yan, "Thermal analysis simulation and test comparison of the water purification tank of railway passenger cars," *Mechanical Engineers*, no. 4, pp. 110–112, 2019.

[6] F. P. Wang and Q. Y. Liu, "Finite element thermal-structure coupling analysis and research on heavy-duty gearbox," *Coal mining machinery*, vol. 32, no. 8, pp. 58–60, 2011.

[7] X. K. Qi, R. P. Shao, and Y. Y. Wang, "Research on thermal-structure coupling analysis of gear box based on UG and ANSYS workbench," *Mechanical rotation*, vol. 39, no. 6, pp. 26–29+35, 2018.

[8] J. Y. Xu, B. Jiang, and F. Sun, "Thermal simulation analysis and optimization design of a new type of ATCA subrack," *Electronic Mechanical Engineering*, vol. 31, no. 6, pp. 61–64, 2015.

[9] H. J. Yoon, M. J. Park, K. B. Shin, and H. S. Na, "Modeling of railway axle box system for thermal analysis," *Applied Mechanics and Materials*, vol. 365-366, pp. 273–276, 2013.

[10] W. T. Yang, W. Cheng, and Z. F. Liu, "Analysis of thermal characteristics of milling head of CNC milling machine in ANSYS," *Mechanical design and manufacturing*, vol. 12, pp. 169-170, 2017.

[11] T. L. Huo, W. J. Ge, and W. Zhao, "Thermal-structure coupling analysis of reducer box," *Mechanical transmission*, vol. 38, no. 4, pp. 133–136, 2014.

[12] M. C. Hsieh, C. K. Yu, and S. T. Wu, "Thermo-mechanical simulative study for 3D vertical stacked IC packages with spacer structures," *Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, vol. 36, no. 6, pp. 47–54, 2010.

[13] W. Shuai and M. Li, "CPU chip heating characteristics and heat dissipation experiments," in *The Tenth National Academic Conference on Engineering Thermophysics in Colleges and Universities*, School of Science, Nanjing University of Science and Technology, Nanjing, China, April 2020.

[14] X. Wang and Y. Shen, "Thermal stress analysis of multi-chip double-sided PCB," *Journal of Nanjing University of Science and Technology (Natural Science Edition)*, vol. 34, no. 2, pp. 18–23, 2010.

Research Article

Classifying Driving Fatigue by Using EEG Signals

Changqing Zeng [iD],[1] **Zhendong Mu,**[2] **and Qingjun Wang** [iD][3,4]

[1]*School of Software, Nanchang University, Nanchang 330047, Jiangxi, China*
[2]*The Center of Collaboration and Innovation, Jiangxi University of Technology, Nanchang 330098, Jiangxi, China*
[3]*College of Economics and Management, Shenyang Aerospace University, Shenyang 110136, Liaoning, China*
[4]*Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*

Correspondence should be addressed to Changqing Zeng; zcq@ncu.edu.cn

Fatigue driving is one of the main reasons for the occurrence of traffic accidents. Brain-computer interface, as a human-computer interaction method based on EEG signals, can communicate with the outside world and move freely through brain signals without relying on the peripheral neuromuscular system. In this paper, a simulation driving platform composed of driving simulation equipment and driving simulation software is used to simulate the real driving process. The EEG signals of the subjects are collected through simulated driving, and the EEG of five subjects is selected as the training sample, and the remaining one is the subject. As a test sample, perform feature extraction and classification experiments, select any set of normal signals and fatigue signals recorded in the driving fatigue experiment for data analysis, and then study the classification of driver fatigue levels. Experiments have proved that the PSO-H-ELM algorithm has only about 4% advantage compared with the average accuracy of the KNN algorithm and the SVM algorithm. The gap is not as big as expected, but as a new algorithm, it is applied to the detection of fatigue EEG. The two traditional algorithms are indeed more suitable. It shows that the driver fatigue level can be judged by detecting EEG, which will provide a basis for the development of on-board, real-time driving fatigue alarm devices. It will lay the foundation for traffic management departments to intervene in driving fatigue reasonably and provide a reliable basis for minimizing traffic accidents.

1. Introduction

The fast-paced life of modern society not only promotes the rapid development of society, but also increases the work pressure of workers. Long-term high-intensity work can easily cause workers' mental fatigue. Mental fatigue will not only reduce work efficiency, but also be harmful to health and even cause accidents. In recent years, fatigue driving has become one of the main causes of traffic accidents. Truck drivers who travel long distances and drivers who work overtime are often prone to fatigue or drowsiness and some even fall asleep while driving. Therefore, fatigue detection has become an important topic in the field of driving safety. Scientists try to use this, to try to reduce the frequency and severity of traffic accidents, and formulate relevant regulations.

At present, there are still many problems in the research on driving fatigue detection technology. For example, the method of detecting the physical reaction of the driver has a high false alarm rate; the method of detecting vehicle parameters requires clear markings on the center line of the road and good light; and the method of measuring the pupil diameter is difficult to implement. In recent years, a large number of scientific researchers have invested in the research of noncontact collection of ECG and EEG and have made great breakthroughs, making the detection of fatigue through biological signals extremely attractive. It provides a basis for further research and development of on-board, real-time driving fatigue alarm devices, and provides a reliable basis for traffic management departments to scientifically and reasonably intervene in driving fatigue and minimize man-made traffic accidents.

At present, the mental fatigue detection at home and abroad mainly selects the following three characteristics:

(1) Behavioral characteristics: behavioral characteristics include changes in lateral lane positions and variability in vehicle heading differences, which are one of the commonly used driving indicators to identify fatigue through vehicle behavior. This method is easy to understand, but the standards for these features have not yet been formed, and it is not easy to achieve stable and effective recognition results.

(2) Facial expression characteristics: facial expression characteristics include the degree of eye closure and the frequency of nodding, etc.

Although this method is very convenient to use, the detection of these features is affected by various factors, such as image angle and image brightness. These limitations reduce the overall recognition accuracy of this method and limit its application scenarios.

(3) Physiological characteristics: physiological characteristics include ECG, heart rate, electromyogram, and EEG. Different from behavioral features and facial features, these features, as the mapping of the physiological changes of the human body, objectively reflect the current state and environment of the human body. Therefore, driving fatigue detection based on physiological features is a very popular application direction. And among these fatigue detection methods based on physiological features, detecting EEG signals is considered the most direct, effective, and promising method to detect driving fatigue.

Fatigue refers to a complex physiological phenomenon in which the physiological and psychological functions of the human body are out of balance due to continuous or high-intensity physical or mental work. Driving fatigue is the driver's excessive consumption of continuous driving or other physical labor for a long time, or lack of sleep, which leads to drowsiness, sluggishness, and weakness in the limbs. phenomenon of the situation. When a driver is fatigued, many features will change, including visible features such as driving posture, blink rate, eye closure time, nodding action, reaction speed, facial expressions, and invisible features such as EEG, ECG signal, EMG signal, skin temperature, skin resistance, respiratory rate, etc. The main manifestations are drowsiness and weakness of the driver, difficulty in eliminating drowsiness, decreased vision, gradually narrowing of the field of vision, inattention, decreased judgment, decreased thinking ability, weak limbs, decreased sensory organ function, unstable driving movements, delayed judgment, Rhythm disorders, knee fatigue, neck stiffness, backache and leg pain, loss of self-control, resulting in a series of uncomfortable symptoms such as being easily excited, irritable or driving fast.

When the driver is slightly fatigued, the shifting is not timely and accurate. In moderate fatigue, the operation is sluggish and sometimes even forgets to operate, In severe fatigue, often subconscious operation or short-term sleep phenomenon, loss of control over the vehicle.

In recent years, how to recognize the fatigue state of a driver has become a hot topic today. Chain R defines an EEG-based driver fatigue test method. The results show that the frequency characteristics of EEG can be well applied to driving fatigue testing, but the overall study lacks data support, and more data is needed to support its conclusions [1]. Yimyam W used 50 recorded EEGs to participate in driving simulation experiments, resulting in two fatigue states: alertness and drowsiness. The experimental results show that as the fatigue degree increases, the two complexity parameters are significantly reduced. The results show that these two nonlinear indicators can be used to characterize driver fatigue. However, the lack of experimental data in his research led to small differences in sample sets, resulting in inaccurate results [2]. Zeng studied the complex, unstable, and nonlinear characteristics of evaluating EEG signals. The results show that: this paper uses the combination of finite element features and AdaBoost classifier to detect driver fatigue based on EEG, ensuring the confidence to explore the internal physiological mechanism and wearable applications. The experiment did not take into account the deviation of the subject's gender and age parameters in the experiment [3].

This paper first extracts more representative EEG signal features by designing a more appropriate feature extraction method, and then selects a suitable classifier to perform optimized to achieve a better classification effect, and classify the signal features extracted by the new feature extraction method through the optimized classifier to achieve the optimal classification effect of the experiment. When studying driving fatigue, this article continuously analyzes the EEG data of driving for 2 hours, trying to find the time when the fatigue occurs. This is something that scholars who have studied driving fatigue have never done before. Most of them only select data from fatigue and normal states to analyze the differences.

2. EEG Signals and the Classification of Driver Fatigue

2.1. Composition and Classification of Brain-Computer Interface.
The exploration of the brain's operating mechanism is not only of great significance in medicine, but also of great help to the development of artificial intelligence. Brain-computer interface is a new technology developed for this purpose, and its application in scientific research is becoming increasingly extensive.

2.1.1. System Composition of Brain-Computer Interface.
The core of brain-computer interface (BCI) technology is to directly collect the EEG of the cerebral cortex through the electrodes and transmit it to the computer via the circuit for signal analysis, and then send the corresponding control commands to the external equipment to achieve a complete EEG control path, instead of passing through normal neuromuscular pathways such as peripheral nerves and

muscle tissues [4, 5]. The main working principle of the brain-computer interface system is: nerve cells generate local potentials when they are active. These potentials can be collected and analyzed in the cortex to analyze the changes in the brain's thinking activity at this time. These changes can be recorded and amplified by the electrode cap. After processing, the collected EEG extraction features are analyzed through information analysis, and the corresponding control instructions are made according to the features to complete the control of the external equipment. The pattern recognition part of BCI is mainly to send the EEG characteristic signal to the classifier and to distinguish the classification space to which the characteristic signal belongs by searching for a suitable decision plane. The result of the classifier is directly related to whether the feature is correctly identified, and therefore directly affects the effect and performance of the BCI. Different classifiers have different advantages in different application scenarios because of their different algorithm principles, so the choice of classifier is essential [6, 7].

A general brain-computer interface system consists of the following five main steps: EEG signal acquisition, signal preprocessing, feature extraction, pattern recognition, and sending external control commands.

The signal acquisition part of BCI is the connection between the system and the brain. This part mainly collects the electrical signals of the cerebral cortex through the electrode cap, then amplifies the electrical signals, and converts the amplified EEG signals into digital signals through A/D conversion. The signal acquisition part of BCI mainly involves the selection of signal types and the number of channels and positions. Since the collected EEG signals will be interfered by EEG, ECG, power, frequency interference, and some high-frequency noise signals, these signals will affect the follow-up, the reliability and accuracy of the processing part. Therefore, the preprocessing part of BCI will use denoising algorithm or filter to remove the noise signal doped in the signal and improve the signal-to-noise ratio of the EEG signal.

EEG signals are used by most BCI systems due to their high temporal resolution. BCI has a wide range of applications in the field of artificial intelligence because of its cheap equipment and no harm to the human body. The BCI system has a variety of classification methods from different perspectives:

Lossy and lossless BCI systems, the acquisition methods of the lossy BCI system can be divided into implantable and semi-implantable. The feature of the implantable BCI system is that the electrodes are implanted into the brain tissue in the cranial cavity, so that the collected brain signals are more accurate; the feature of the semi-implantable BCI system is that the electrodes are attached to the cerebral cortex in the cranial cavity, and the brain Cortical potential changes can be accurately recorded by the electrodes.

Induced and spontaneous BCI systems, the evoked BCI system has many kinds of elicitation methods, which can be divided into exogenous stimulation and endogenous stimulation according to the different stimulation sources. Potential changes induced by exogenous stimuli are called exogenous evoked potentials, including visual stimulus evoked potentials, auditory stimulus evoked potentials, and tactile stimulus evoked potentials. Endogenous stimuli are related to cognitive processes in the brain, and endogenous evoked potentials are brain potentials evoked by specific, multiple or diverse stimuli.

2.1.2. Application Fields of Brain-Computer System

(1) Rehabilitation Field. The BCI system has the most urgent needs in the field of rehabilitation. There are many patients with physical disabilities throughout the country. For them, the BCI system provides them with a new channel for communication with the outside world, which can significantly improve their quality of life and even live a normal life.

(2) Military Field. Although the current research on related technologies has not reached the stage of maturity in movies, the concept of using thinking activities in military operations is the consensus of all countries in the world, many countries have spent a lot in this field, and movies have also been realized.

(3) Leisure and Entertainment Field. Leisure and entertainment have always been an indispensable part of human life, so the research direction of brain-computer interface is also biased towards this [8]. Now, ordinary keyboard and mouse devices can no longer keep up with people's response speed and computer processing speed. People urgently need a new control method to control computers more conveniently for entertainment activities. Simple such as browsing the web, typing on the Internet, designing drawings, complex multiline operations, controlling games, These applications will greatly change the way of life and entertainment in the future, have large market potential, and attract countless teams to develop related technologies.

(4) The Field of Brain Science Research. The research of the BCI system originally originated from the study of brain cognitive science, such as the use of brain electricity to detect polygraphs, the study of mental activity based on brain electricity, and this article involved in driving fatigue detection based on BCI system [9]. The research of the BCI system is based on the research of brain science, and the research of the BCI system in turn provides help to the research of brain cognitive science.

2.2. EEG Signal.

Generally speaking, EEG refers to the electrophysiological phenomenon produced by the activity of brain nerve cells in the cerebral cortex, which is a kind of bioelectric signal. Since the human body is a complex system, the electrical signals collected through the scalp include not only EEG, but also physiological signals such as my electricity and ocular electrical signals, as well as noise interference such as instrument power frequency interference, so the effect of directly processing the collected EEG is not ideal. It is necessary to preprocess the EEG and then conduct research and analysis [10, 11].

People will experience various subjective feelings when driving fatigue, such as dizziness, blurred vision, difficulty concentrating, feeling irritable, feeling sleepy, muscle weakness, reference basis. When driving fatigue, the subjects will have many weakened functions controlled by the brain. Different parameters of the EEG signal represent different activities of the human brain and different mental states. However, the accurate extraction of these parameters requires good technical means suitable for EEG signals.

2.2.1. The Structure and Function of the Brain.
The brain is the most complex and mysterious organ in the human body, and there are many mysteries that have not yet been completely solved. The brain is divided into two parts: the left brain and the right brain. According to the existing medical data, the left and right hemispheres are responsible for different functions such as thinking and imagination, and are connected to each other through the corpus callosum. The left and right hemispheres are subdivided into four parts: frontal lobe, parietal lobe, temporal lobe, and occipital lobe. These areas control different functions. According to some latest brain science research data, the various functions of the brain are not only completed by the abovementioned parts alone, but are the result of the collaborative processing of various parts of the brain [12–14]. Even after some parts are damaged, the brain can coordinate the other parts to complete the functions of the damaged part of the brain. The cerebral cortex is the general term for the gray matter of the brain's surface layer, and its thickness is about 1.5–5.0 mm. The gray matter of the cerebral cortex in different regions corresponds to different functions, as shown in Figure 1.

According to scientific research data, at this stage, the cerebral cortex contains tens of billions of nerve cells. With this large group of nerve cells, the brain can achieve very complex functions. For driving fatigue detection based on EEG, it is first necessary to understand the relationship between mental fatigue and the state of different areas of the cerebral cortex and time, to accurately judge the state of mental fatigue.

2.2.2. The Generation Mechanism of EEG Signals.
Generally speaking, EEG refers to the local electrical signal produced by the activity of nerve cells in the cerebral cortex, which is produced by the electrophysiological activity of a large number of nerve cells in the relevant areas of the brain. According to the research data, the dendrites of nerve cells will generate an action potential after receiving stimulation. The potential will be transmitted along the axon to the axon terminal, and the potential will be transmitted to the next nerve cell through the axon terminal. The electrical signal generated by these large nerve cells transmitting potential is EEG [15, 16].

2.2.3. Characteristics and classification of EEG signals

(1) Randomness: because the cerebral cortex is not completely insulated, the electrical signals of different areas will influence and interfere with each other, sometimes accompanied by unconscious light discharge behaviors. Therefore, EEG has significant randomness, which humans cannot completely present. Analyzing the trend of EEG can only roughly analyze the characteristics of EEG.
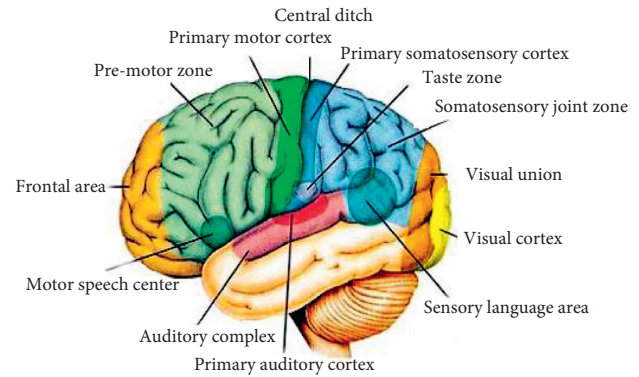
(2) Nonstationary: the generation of brain signals is affected by many factors, so it has the characteristics of nonstationary. Using this feature of EEG, the amplitude and frequency of EEG can be compared to distinguish whether the brain is in a state of fatigue.

What we usually call an EEG refers to the brain scalp EEG, which is actually a graph of the relationship between the scalp potential difference and time. Brain waves are the comprehensive external manifestations of the overall activity of brain nerve cells, including ion exchange and metabolism. The EEG signal is an electrical phenomenon displayed on the surface of the cerebral cortex or scalp by the physiological activity of brain nerve cells. This electrical phenomenon accompanies a person's life from the beginning to end. Once a person dies, the electrical phenomenon disappears. Therefore, the EEG signal is a signal obtained by extracting the spontaneous and rhythmic physiological activity of the brain cell population on the surface of the scalp and amplifying it by a million times. EEG signals can be divided into spontaneous EEG and evoked EEG. Spontaneous EEG refers to changes in EEG without external stimulation. Induced EEG is a change in brain potential caused by artificially stimulating human sensory organs.

(3) Nonlinearity: the generation of EEG is based on the multiple coupling of human physiological state and life phenomena, so it has chaotic characteristics and nonlinearity. Simply using linear analysis of EEG cannot fully reflect the characteristics of EEG. It is necessary to use nonlinear analysis EEG to better analyze the internal information of EEG [17].

(4) Low amplitude, low frequency, and low signal-to-noise ratio: EEG signals have the characteristics of low amplitude, low frequency, and low signal-to-noise ratio. Since the EEG signal is the bioelectricity

Figure 1: Distribution map of cerebral cortex function.

generated by the activity of nerve cells, the amplitude of the EEG signal is relatively low only at the micro- and millivolt level, generally not exceeding $100\,\mu V$. The frequency of EEG signals is relatively low, generally between 0.5 and 100 Hz. It is precisely because of the characteristics of low amplitude and low frequency of EEG signals that it is easily interfered by other external signals and has the characteristics of low signal-to-noise ratio.

EEG signal is a manifestation of brain nerve activity and an important signal reflecting the state of brain activity. Its amplitude range is one and its frequency range is one. The rhythm of EEG signals changes with the change of a person's mental state and a series of transients such as sharp waves or spikes in epilepsy patients also appear. Usually, the collected EEG signals are very weak, and there is a lot of background noise and strong interference signals, such as power, frequency interference, ECG interference, breathing interference, swallowing, eye blinking, which increase the difficulty of analysis. However, because the method is a noninvasive method and contains a lot of physiological and pathological information, it still attracts the interest of many researchers. They focus on the time-varying nonstationary characteristics of EEG signals. It has continuously proposed analytical methods that can effectively characterize different physiological states of the brain [18, 19].

2.3. Feature Extraction Based on Frequency Band Power Spectrum.

When the human brain is active, it produces five frequency bands: alpha, gamma, theta, delta, and sigma. When the power spectrum peaks of the alpha wave and the beta wave are obvious, which are the main bands of the EEG, it indicates that the subject is awake; when the subject feels tired or even sleeps the peaks of theta wave and delta rise significantly [20, 21]. Using the STFT algorithm to extract the power spectrum is a power spectral density function (PSD) feature as follows:

$$STFT\{x[n]\}(m, w_k) = \sum_{n=1}^{N} x[n]w[n-m]e^{-jw_k n}, \quad (1)$$

where $w_k = (2\pi k/N)$, $k = 0, 1, \ldots, N-1$ is the angular frequency, m is discrete and ? is continuous, its mathematical expression is as follows:

$$w[n] = 0.5\left[1 - \cos\left(\frac{2\pi n}{(N-1)}\right)\right] = \sin^2\left(\frac{\pi n}{(N-1)}\right). \quad (2)$$

The energy spectrum at different frequencies of different EEG frequency bands is calculated as follows:

$$E(w_k) = X(m, w_k)X^*(m, w_k),$$

$$PSD(w_k) = \frac{1}{N}E(w_k), \quad k = 0, 1, \ldots, N-1. \quad (3)$$

The energy spectrum of each EEG band can then be defined as:

$$E(w) = \sum_{k=1}^{n} E(w_k), \quad (4)$$

where $x[n] = \{x_1, x_2, \ldots, x_n\}$ represents all frequencies in a frequency band.

2.4. Feature Extraction Method Based on EMD and Energy Spectrum

2.4.1. Empirical Mode Decomposition. Empirical Mode Decomposition (EMD) decomposes a given signal $x(t)$ into a series of Intrinsic Mode Function (IMF) components from high to low. EMD can be regarded as a type of wavelet decomposition. Its subbands are established according to the decomposition of each component of the signal. Each IMF component generated represents the details of the signal on a certain scale and frequency band [22, 23]. The steps of EMD decomposition are as follows:

(1) Identify all maximum points in the EEG and fit the upper envelope $e_{up}(t)$ of the signal, and identify all minimum points in the EEG and fit the lower envelope $e_{low}(t)$ of the signal.

(2) Calculate the average value $m_1(t)$ according to the synthesized upper and lower envelopes, the formula is

$$m_1(t) = \frac{e_{up}(t) + e_{low}(t)}{2}. \quad (5)$$

(3) Subtract $x(t)$ from $m_1(t)$ to obtain $h_1(t)$, and use the obtained $h_1(t)$ as the new EEG $x(t)$. Repeat step 1). After k screening, $h_1(t)$ at this time is the first IMF component of the signal, denoted as $c_1(t) = h_1(t)$.

(4) Repeat the above steps with the difference $r_1(t) = x(t) - c_1(t)$ between $c_1(t)$ and $x(t)$ as a new signal to obtain a series of IMF components. When the obtained IMF component or residual signal is less than a preset value, the empirical mode decomposition step is completed.

2.4.2. Principle of Energy Spectrum. Energy spectrum (Energy Spectral Density, ESD), arrangement of the harmonic energy contained in the signal, is called the energy spectrum, which describes the energy contained in each frequency component in the signal. The formula of the energy spectrum is as follows:

$$\Phi(w) = \left|\frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} f_n e^{-jw_k n}\right|$$

$$= \frac{F(w)F^*(w)}{2\pi}. \quad (6)$$

In the formula, f_n is the signal sequence. Divide the observation data x(t) of length N into K segments of data of length M, that is, $N=KM$. Estimate the third moment of each piece of data $x_i(t)$ ($i = 1, 2, \ldots, k$):

$$\widehat{r}(i)(m,n) = \frac{1}{M} \sum_{k=s_1}^{s_2} x_i(t) x_i(t+m) x_i(t+n), \ i = 1, 2, \ldots, k. \quad (7)$$

Among them $S_1 = \max(1, 1-m, 1-n)$, $S_2 = \min(M, M-m, M-n)$. Find the overall average of the third-order moment estimates of these K segments of the data, and find the third-order cumulant estimate of x(t):

$$\widehat{r}(m,n) = \frac{1}{K} \sum_{i=1}^{k} \widehat{r}^{(i)}(m,n). \quad (8)$$

By estimating the third moment $\widehat{\beta}$ of the non-Gaussian white noise sequence w(k), the parameters of the AR model are obtained:

$$\widehat{r} \cdot a = \widehat{b},$$

$$\widehat{r} = \begin{bmatrix} \widehat{r}(0,0) & \widehat{r}(1,1) & \cdots & \widehat{r}(p,p) \\ \widehat{r}(-1,-1) & \widehat{r}(0,0) & \cdots & \widehat{r}(p-1,p-1) \\ \vdots & \vdots & \widehat{r}(0,0) & \vdots \\ \widehat{r}(-p,-p) & \widehat{r}(-p+1,-p+1) & \cdots & \widehat{r}(0,0) \end{bmatrix},$$

$$\widehat{b} = [\widehat{\beta}, 0, 0, \ldots, 0]^T. \quad (9)$$

A is the AR model parameter to be estimated for p. Use p model parameters to estimate the energy spectrum:

$$B(w_1, w_2) = \widehat{\beta}\widehat{I}(w_1)\widehat{I}(w_2)\widehat{I}^*(w_1 + w_2). \quad (10)$$

Among them,

$$\widehat{I}(w) = \left[1 + \sum_{n=1}^{p} a_n \exp(-jwn)\right]^{-1}, \ |w| \le \pi. \quad (11)$$

2.4.3. Feature Extraction Based on EMD and Energy Spectrum.

Traditional EMD-based feature extraction methods usually contain a large number of time-domain signals, which is very helpful for extracting the main components of the signal, but it is not ideal for the extraction of fatigue signals. The energy spectral density function can extract the frequency domain information of each IMF component and extract the time-frequency information of EEG by combining it with EMD decomposition. The flow chart of the algorithm is shown in Figure 2.

It can be seen from Figure 2 that after the original EEG has been preprocessed, the irrelevant noise signals have been roughly filtered out. The first three layers of IMF components are selected for processing, and after feature extraction, 96 energy spectrum features (32 channels × 3 layers of IMF) are obtained. Compared with the power spectrum feature extraction method, it has fewer features and requires less calculation and calculation time for classification.

3. Experimental Design of Driver Fatigue Classification Based on EEG Signal

3.1. Simulated Driving Experiment Design.
The EEG signals in this experiment are all collected by laboratory simulation driving. In this article, the real driving process is simulated by using a driving simulation platform composed of driving simulation equipment and driving simulation software. The driving simulator platform is an advanced simulator system that can simulate real driving conditions while providing dynamic displays of vehicles and surrounding traffic. The driving simulation platform itself consists of a fixed car, steering wheel, brake pedal, and accelerator pedal, a large screen, a high-performance computer, driving simulation software, and a multifunctional data acquisition board. The platform can dynamically record EEG during the mission and calculate the driving and running status of the vehicle.

The EEG acquisition equipment used in this experiment is an EEG signal acquisition and analysis system composed of Brain Amp, DC amplifier of German Brain Products, and corresponding software. The system includes E-Prime stimulation software, actiCAP electrode cap, Brain Amp amplifier, Brain Vision Recorder recording Software, and Brain Vision professional Analyzer analysis software, the physical object of the equipment is shown in Figure 3.

3.2. Experimental Objects and Methods.
In this experiment, the EEG data of 5 participants was selected as the training set, and the remaining 1 data was used as the test sample for feature extraction and classification experiments, pick out two sets of data from the data, one is normal signal data, and the other is fatigue signal data, and analyze them. We asked the participants to sleep only 4 hours as fatigue state data the night before data collection, and 8 hours sleep as waking state data, to obtain EEG data in the two states. The participants' simulated driving time was set at 8 pm the next day, during which EEG recording was started at 8:10, and signal recording was stopped at 8:30 pm.

The EEG acquisition experiment took 20 minutes. During the experiment, the subject was asked to sit on a chair and simulate the driving process on the driving simulation platform. During the experiment, the EEG was recorded and used with a 32-channel electrode cap according to the international 10–20 lead system. Brain Vision Recorder records EEG with a sampling rate of 1000 Hz. The temperature of the laboratory is maintained at 22°C, which provides the subjects with a comfortable and stable experimental environment. Figure 4 shows the actual scenario of the experiment.

3.3. EEG Signal Collection.
Due to the different functional composition of various regions of the brain, the EEG signal bands appearing in different regions will also change. In addition, EEG signals have the characteristics of low amplitude and low frequency, which are easily disturbed by noise signals. Therefore, how to choose appropriate electrode selection location is essential. Using Brain Vision Recorder recording software to record the EEG during driving, it can be seen that most of the electrodes are in good contact with the scalp, and all signals meet the requirements of EEG collection. The EEG of fatigue and normal driving were collected, respectively. Each subject collected 20 minutes of normal signal and 20 minutes of fatigue
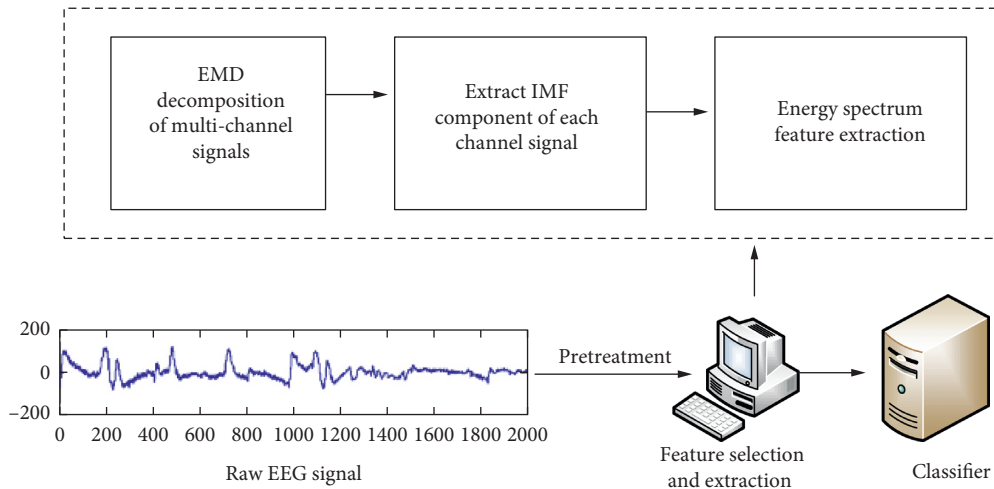
Figure 2: Flow chart of feature extraction algorithm.
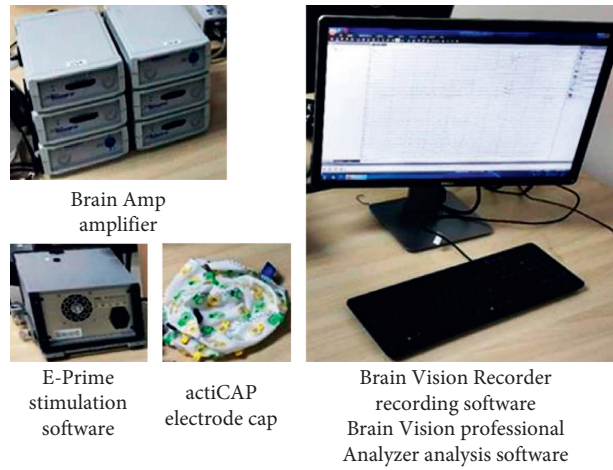
Figure 3: EEG acquisition system(https://image.cnki.net/, author authorization is available).
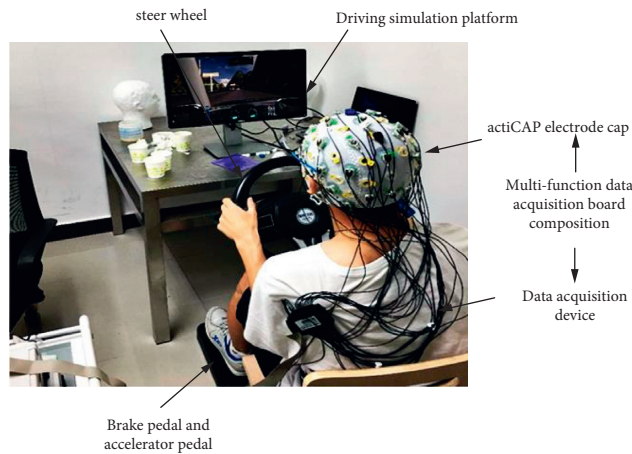
Figure 4: Simulated driving experiment diagram. (https://image.cnki.net/, author authorization is available).

signals. The signals collected every 10 seconds were used as a sample to construct training samples and test samples. In this experiment, six subjects collected a total of 720 groups of normal EEG samples and 720 groups of fatigue EEG samples, and the collected 1440 groups of EEG data samples were subjected to subsequent analysis and processing.

3.4. Frequency Reduction and Filtering of EEG Signals. To study the driving fatigue EEG, it is necessary to analyze the five main EEG bands of alpha, beta, gamma, theta, and delta. The frequency bands of these five signals are roughly between 0.1 and 50 Hz, which belong to low-frequency signals, and the acquisition equipment, sampling frequency is 1000 Hz, which not only carries a large amount of irrelevant high-frequency noise and interferes with data processing, but also a large amount of data processing also causes a lot of burden on the processing and operation of the brain-computer interface device. Therefore, to better carry out the subsequent processing, the collected EEG is first subjected to frequency reduction processing, and the sampling frequency is reduced to 200 Hz, which greatly reduces the amount of data calculation and speeds up the data processing.

3.5. Denoising of EEG Signals

3.5.1. Hard Threshold and Soft Threshold Denoising. Noise shows strong randomness in the wavelet domain and can be considered as Gaussian noise. According to the characteristics of Gaussian distribution, most of the noise coefficients are in the interval $[-3\sigma, 3\sigma]$. Therefore, as long as the threshold is set to $\lambda = 3\sigma$, most of the noise can be removed. Among them, the hard threshold wavelet denoising formula is as follows:

$$\widetilde{d}_{j,k} = \begin{cases} d_{j,k}, & |d_{j,k}| \geq \lambda, \\ 0, & |d_{j,k}| < \lambda. \end{cases} \tag{12}$$

Among them, $d_{j,k}$ is the wavelet coefficient, $\widetilde{d}_{j,k}$ is the estimated value of the wavelet coefficient after hard threshold denoising, and λ is the threshold.

However, directly removing these wavelet coefficients will cause abrupt changes in the wavelet domain, resulting in local jitter in the denoised signal, which will affect the analysis and processing of the signal. Therefore, a soft threshold wavelet denoising method is proposed. Subtract 3σ from the coefficients greater than 3σ, and add 3σ to the coefficients less than -3σ. The formula of the soft threshold wavelet transform denoising method is as follows:

$$\widetilde{d}_{j,k} = \begin{cases} d_{j,k} - \lambda, & d_{j,k} \geq \lambda, \\ d_{j,k} + \lambda, & d_{j,k} \leq -\lambda, \\ 0, & |d_{j,k}| < \lambda. \end{cases} \tag{13}$$

After soft threshold processing, the signal after denoising is smoother than that after hard threshold denoising. At the same time, it will also cause partial loss of detail signal.

3.5.2. Improve Soft Threshold Denoising. The improved soft threshold method combines the advantages of soft and hard threshold methods and can effectively remove low-frequency noise. The formula is as follows:

$$\widetilde{d}_{j,k} = \begin{cases} a\lambda^2 - \lambda + d_{j,k}, & d_{j,k} \geq \lambda, \\ ad_{j,k}^2, & 0 < d_{j,k} < \lambda, \\ -a\lambda^2 + \lambda + d_{j,k}, & d_{j,k} \leq -\lambda, \\ -ad_{j,k}^2, & -\lambda < d_{j,k} < 0, \end{cases} \quad a < \frac{1}{\lambda}. \tag{14}$$

Among them, a is the shape coefficient. The value of a can control the shape of the function in the $d_{j,k} < \lambda$ area, that is, the degree of attenuation. When the value is 0, the formula for improving the soft threshold method is equivalent to the soft threshold method. The selection formula of the threshold is

$$\lambda_j = \frac{\sigma\sqrt{21g(N)}}{\lg(j+1)}, \tag{15}$$

where $b \geq 0$, j is the number of decomposed layers, N is the number of wavelet transform coefficients corresponding to the j layer, λ_j is the wavelet threshold, s is the variance of the noise, and in real life, the way the noise is generated is various, so the value of s is generally estimated according to the first-level wavelet packet decomposition coefficient, the estimation formula is as follows:

$$\widehat{\sigma} = \frac{\sum_{k=1}^{N} |d_{1k}|}{0.6745 * N}. \tag{16}$$

Among them, d_{1k} is the wavelet decomposition coefficient of the first wavelet decomposition layer, and 0.6745 is the adjustment coefficient of Gaussian white noise.

4. Experimental Driver Fatigue Classification Based on EEG Signal

4.1. Compare the Denoising Effects of Different Algorithms. We tested the effect of the improved method, the signal used is a piece of data from the Fp1 channel of subject No. 1. The sampling frequency after frequency reduction is 200 Hz, and it has passed the bandpass filter of [0.1 Hz, 50 Hz]. Processing. The wavelets in the experiment all decompose the original signal in three layers and extract coefficients. The shape coefficient of the soft threshold is 0.01 and the scaling factor is 0.5. The result is shown in Figure 5.

It can be seen from Figure 5 that there are many glitches in the signals of the hard threshold denoising and soft threshold denoising methods, therefore, the signal of the improved method is clearly better than the other three signals. Although the hard threshold denoising method and the soft threshold denoising method have certain effects, the signal effect after the improved soft threshold denoising processing is the best. This shows that the improved soft threshold denoising method can effectively combine the traditional hard threshold and soft threshold. Advantages of the method.
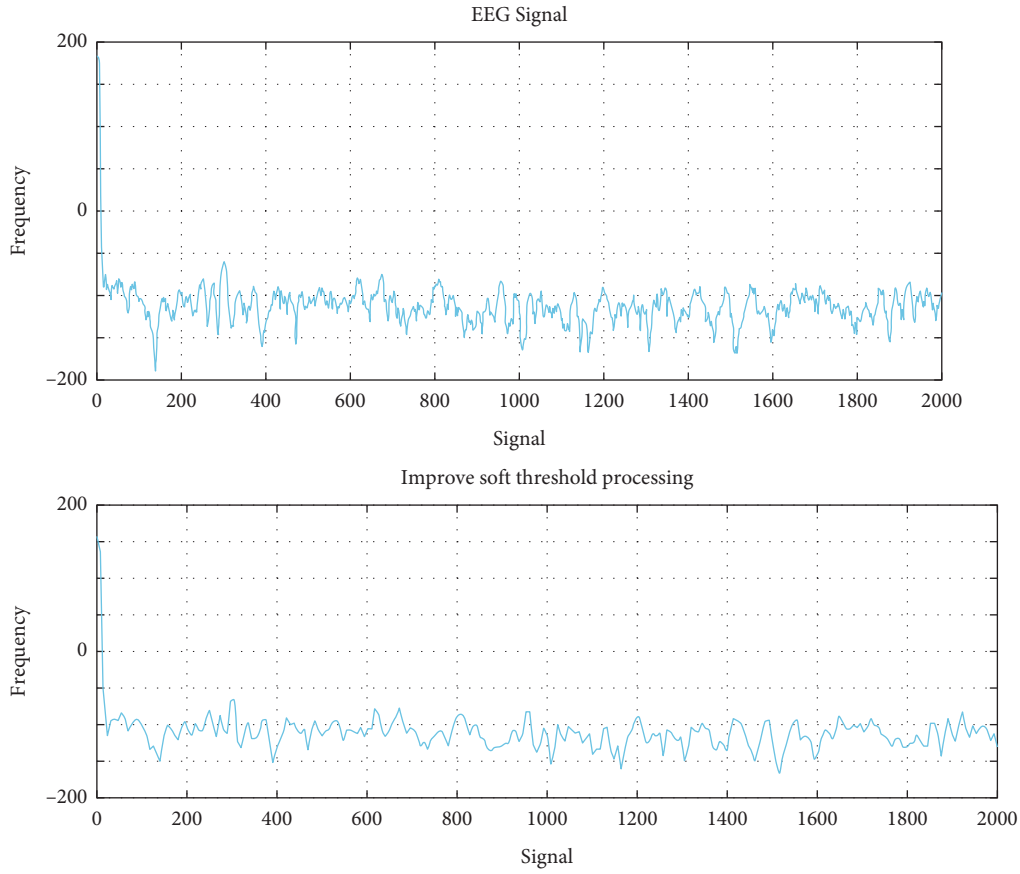
FIGURE 5: Experimental comparison of three wavelet threshold denoising methods.

Want to quantitatively analyze the pros and cons of the three wavelet threshold denoising methods, the signal-to-noise ratio (SNR), and root mean square error (RMSE) are again selected to judge the quality of the denoised EEG. The formula is as follows:

$$SNR = 10 * \log_{10}\left(\frac{\sum_{i=1}^{N} x^2(i)}{\sum_{i=1}^{N} \widetilde{x}^2(i)}\right),$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x(i) - \widetilde{x}(i))^2}.$$

(17)

Among them, N, $x(i)$, $\widetilde{x}(i)$ are the length of the EEG, the original EEG, and the EEG after noise reduction. Randomly select 100 groups of EEG and obtain the signal-to-noise ratio and the root mean square error, respectively, and calculate the average value after the denoising experiment. The experimental results are shown in Table 1.

It can be seen from Table 1 that the improved soft threshold denoising method has the largest SNR, while the RMSE is the smallest, and the denoising effect is the best.

4.2. EM and Energy Spectrum Feature Extraction Results. In this experiment, the EEG data of 5 subjects is selected as the training sample set, and the remaining 1 data is used for feature extraction and classification experiments as the

TABLE 1: SNR and RMSE results after denoising with different threshold methods.

Evaluation index	Threshold method		
	Hard threshold	Soft threshold	Improve soft threshold
SNR	14.49	14.36	18.42
RMSE	276.32	278.96	197.16

sample test. And any set of normal signals and fatigue signals recorded in the driving fatigue experiment were selected for data analysis. In the experiment, the preprocessed signal is selected for feature extraction. After EMD decomposition and energy spectrum feature extraction, the data is compressed to 84 feature points; after power spectrum feature extraction, the EEG data is compressed to 144 feature points. The preprocessed EEG is compared with the signal after the two feature extraction algorithms, and the results are shown in Figures 6–8.

It can be seen from Figure 6 that before the feature extraction, the signals of the normal signal and the fatigue signal are entangled, which is very difficult to judge. It can be seen from Figure 7 that after using the power spectrum to perform feature extraction on the signal, it can be seen that the low-frequency range of the EEG in the fatigue state fluctuates more, and it can be easily distinguished from the normal EEG. It can be seen from Figure 8 that after EMD decomposition combined with energy spectrum for feature
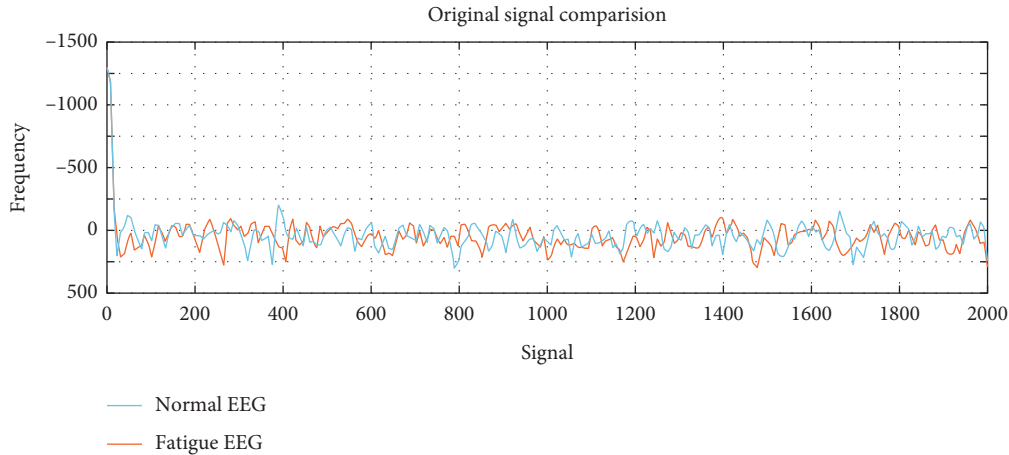
FIGURE 6: Original signal comparison.
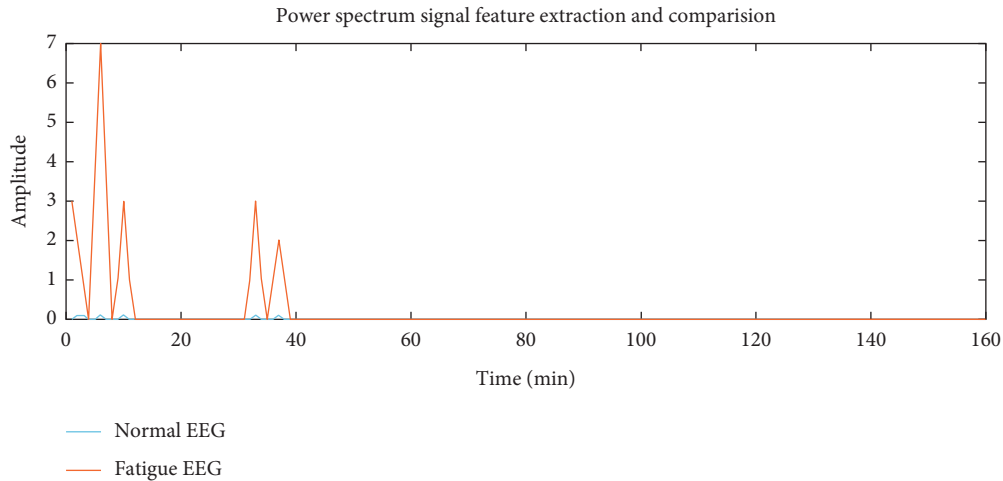
FIGURE 7: Power spectrum signal feature extraction and comparison.

extraction, the energy difference feature in the low frequency range is more obvious than the power spectrum feature. Therefore, it can be assumed that the feature extraction method combined with EMD decomposition and energy spectrum is more suitable than the power spectrum. EEG recognition of driving fatigue.

Want to understand the pros and cons of the two feature extraction algorithms, and analyze the optimization effect of the PSO-H-ELM algorithm, so I select the EEG data of 5 participants as the training set, and the EEG data of the remaining participants as the test set. The experimental result is the operation of different subjects as the test set, as shown in Table 2.

It can be seen from Table 2 that although the test results of different subjects are biased, it can be found from the average accuracy rate that the features extracted using EMD decomposition combined with energy spectrum are using SVM, KNN and PSO-H-ELM three classifiers In the case, the classification effect is better than using the power spectrum feature extraction method. At the same time, the results show that the algorithm can achieve the best classification

effect after using the algorithm of EMD decomposition combined with energy spectrum for feature extraction.

4.3. Classification Performance of Different Algorithms.
Select 240 sample data of 1 participant as the test set, and 1200 sample data of the remaining 5 participants as the training set. The purpose of this arrangement is to avoid confusion between training and test data. The results are shown in Table 3.

It can be seen from Table 3 that the accuracy of 72.32%, 71.47%, 89.25%, 93.17%, 73.85%, and 70.79% were obtained by using KNN. The accuracy rates of support vector machines are 77.25%, 74.67%, 88.32%, 94.83%, 63.29%, and 73.33%, respectively. The accuracy rates of ELM are 70.85%, 66.83%, 86.25%, 88.77%, 58.25% and 67.75%, respectively. The accuracy rates of H-ELM are 78.17%, 81.08%, 89.24%, 89.38%, 71.78%, and 74.24%, respectively. Finally, the accuracy of 78.91%, 83.85%, 91.25%, 91.78%, 73.33%, and 76.75% were obtained using PSO–H-ELM. The accuracy of the algorithm changes greatly when different subjects are
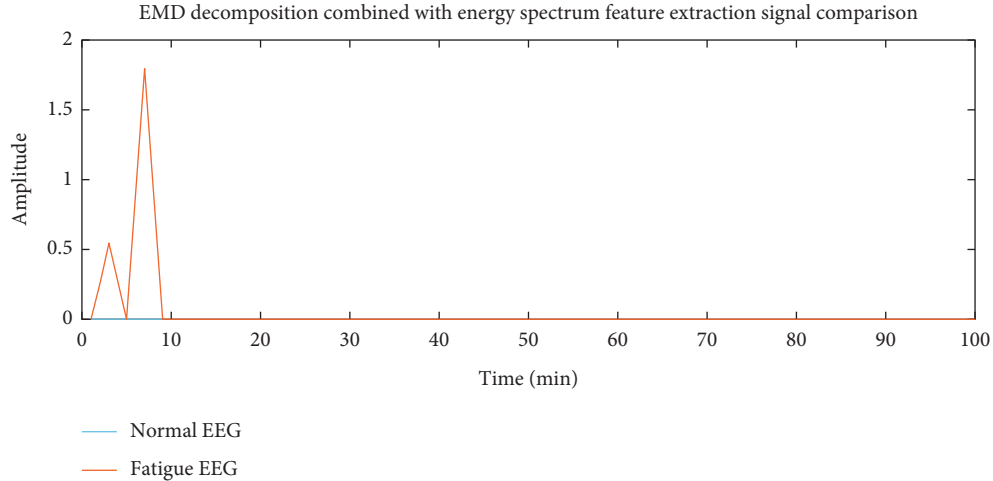
FIGURE 8: EMD decomposition combined with energy spectrum feature extraction signal comparison.

TABLE 2: Different feature extraction algorithms using SVM, KNN, and PSO-H-ELM classifiers to classify the accuracy (%).

Feature extraction algorithm	Classification algorithm	1	2	3	4	5	6	Average accuracy
Power spectral density	SVM	81.43	83.59	94.38	89.76	70.79	74.13	83.33
	KNN	77.39	77.57	97.33	87.57	88.58	89.32	86.28
	PSO–H-ELM	81.43	80.76	97.73	97.06	85.75	85.26	88.83
EMD	SVM	90.75	85.38	88.25	100.00	97.29	95.67	93.62
	KNN	77.25	83.33	88.71	100.00	97.25	95.67	90.94
	PSO–H-ELM	89.60	88.85	94.25	69.57	99.85	98.23	94.24

TABLE 3: Classification accuracy using different training samples (%).

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
KNN	72.32	71.47	89.25	93.17	73.85	70.79
SVM	77.25	74.67	88.32	94.83	63.29	73.33
ELM	70.85	66.83	86.25	88.77	58.25	67.75
HELM	78.17	81.08	89.24	89.38	71.78	74.24
PSO-HELM	78.91	83.85	91.25	91.78	73.33	76.75

used as training samples. Therefore, it is difficult to directly determine from Table 3 which classification algorithm has the best effect. To make the comparison of the experimental results more intuitive, the average classification accuracy rate of 5 classifiers is calculated, as shown in Table 4.

It can be seen from Table 4 that the average accuracy of the PSO–H-ELM algorithm is the highest, indicating that the proposed PSO–H-ELM algorithm is superior to other algorithms in driving fatigue detection. By comparing the performance of H-ELM and the proposed PSO–H-ELM in different situations, it can be found that the change of training data does not affect the optimization of the H-ELM algorithm by the PSO algorithm.

To confirm the results of the experiment, paired t-tests are used to test the accuracy of PSO-H-ELM to see if it has obvious advantages compared to other classifiers. The results of the statistical test are shown in Table 5.

It can be seen from Table 5 that the PSO–H-ELM algorithm has only about 4% advantage compared with the average accuracy of the KNN algorithm and the SVM algorithm. The gap is not as large as expected, but it is used as a

new algorithm for fatigue EEG. The detection is indeed more suitable compared to the two traditional algorithms. When comparing the PSO–H-ELM algorithm to the ELM algorithm and the H-ELM algorithm, it can be found that the advantages of the PSO–H-ELM algorithm are very obvious and stable. PSO–H-ELM is an optimized algorithm for H-ELM. The classification and recognition of EEG is much more accurate than that of H-ELM algorithm. Compared with traditional mainstream algorithms, the PSO–H-ELM algorithm is more suitable for the classification of driving fatigue EEG.

5. Discussion

(1) Living environment: the place of residence is too far from the work place; there are too many housework or the husband and wife are not in harmony; the mental burden is heavy; (2) Sleep quality: going to be too late, too little sleep time; poor sleep effect; noisy sleep environment cannot guarantee sleep quality. (3) In-car environment: poor air quality, poor ventilation; too high or too low temperature;

TABLE 4: The average classification accuracy of the algorithm (%).

	KNN	SVN	ELM	HELM	PSO-HELM
Accuracy	79.31	79.31	74.08	81.67	83.12

TABLE 5: The correct rate of each classifier t-test test.

	Pairing bias					t	df	Sig.(2-tailed)
	Mean	Standard deviation	Mean error	95% confidence interval				
				Lowest	Highest			
P–KNN	3.820	4.985	1.949	-1.192	8.832	1.949	5	0.103
P-SVM	3.820	4.923	2.006	-1.338	8.821	1.834	5	0.119
P-ELM	9.833	4.576	1.867	4.791	14.404	5.133	5	0.004
P-HELM	1.453	0.683	0.279	0.741	2.175	5.127	5	0.001

bad noise and vibration; improper seat adjustment; tense relationship with fellow passengers. (4) Environment outside the car: driving in the afternoon, evening, early morning, and late at night; poor road conditions; good road conditions, single situation; driving in windy, rainy, foggy, snowy weather; poor traffic environment or crowded traffic conditions. (5) Operating conditions: long-term, long-distance driving; too fast or too slow; too limited time to reach the destination. (6) Physical condition: poor physical strength and endurance; decreased visual and hearing ability; weak physical strength or suffering from certain chronic diseases; taking drugs that are contraindicated in driving vehicles; special physiological periods for women (menstrual period, pregnancy). (7) Driving experience: low technical level and unfamiliar operation; short driving time and little experience; poor safety awareness.

Driving at high speed can easily make the driver nervous. Driving at a slightly lower speed than normal without affecting traffic can reduce mental stress and fatigue. At the same time, we try to reduce overtaking when driving, reduce emergency braking, and other actions, which can also reduce fatigue. When you feel sleepy during driving, you can also open the window to ventilate (to prevent endogenous oxygen deficiency), play different styles of music, chew gum, apply wind oil essence, and body spirit dew is more effective to prevent fatigue driving. A must-have, when you are sleepy, apply on the Renzhong acupoint, and take a deep breath to keep your brain awake and reduce fatigue. Under the premise of not affecting safety, you can stretch your fingers, shake your arms, twist your neck, and look as far away as possible. Always staring at the center line will cause visual fatigue.

If you drive a motor vehicle for more than 4 hours continuously, you should stop and rest, and the rest time should not be less than 20 minutes. Passenger vehicles, heavy-duty trucks, and semitrailer tractors used for highway operation shall install and use driving recorders that meet national standards. Traffic police can check the driving speed, continuous driving time, and other driving status information of the motor vehicle.

6. Conclusions

In daily life, the driver cannot always drive the vehicle in a stable state. When the driver finishes a day's work, his brain power becomes exhausted, and it is very easy to experience fatigue when driving a vehicle. To avoid the occurrence of fatigue driving, this article uses the stable and reliable characteristics of EEG, collects EEG through the brain-computer interface, and processes and analyzes it, to achieve the purpose of judging and early warning of fatigue driving. This paper implements a feature extraction method that combines empirical mode decomposition and energy spectrum: the data is decomposed into multiple IMF components through empirical mode decomposition, and then the energy spectrum is used to extract features for each IMF component, and these components are combined at the feature points to construct the feature vector of EEG. The experiment uses the new method and the power spectrum method to extract the features of EEG and uses a variety of classifiers for classification processing. Compare the classification experiment results of the two methods, and the better feature extraction effect of the new method is verified. This research still has certain limitations. For example, there are still some differences between simulated driving and actual driving, the simulated driving time is not long enough, the time when drivers of different physiques are fatigued under different road conditions, and the characteristics of fatigue will be different. There are still many problems in the development of driving fatigue research, such as real-time problems, contact problems, and cost-effectiveness of detection devices.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Chai, G. R. Naik, T. N. Nguyen et al., "Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 715–724, 2017.

[2] W. Yimyam and M. Ketcham, "Driver fatigue F monitoring algorithm using support vector machine and first-order neighborhood mean-shift," *Far East Journal of Electronics and Communications*, vol. 18, no. 5, pp. 773–792, 2018.

[3] H. Zeng, C. Yang, G. Dai, F. Qin, J. Zhang, and W. Kong, "EEG classification of driver mental states by deep learning," *Cognitive Neurodynamics*, vol. 12, no. 6, pp. 597–606, 2018.

[4] Z. Mu, J. Hu, and J. Min, "Driver fatigue detection system using electroencephalography signals based on combined entropy features," *Applied Sciences*, vol. 7, no. 2, pp. 150–161, 2017.

[5] M. Hadra and I. Abdelrahman, "Automatic EEG-based alertness classification using sparse representation and dictionary learning," *Journal of Biomedical Engineering and Medical Imaging*, vol. 7, no. 5, pp. 19–28, 2020.

[6] W. Yimyam and M. Ketcham, "The system for driver fatigue monitoring using decision tree via wireless sensor network for intelligent transport system," *International Journal of Online Engineering (iJOE)*, vol. 14, no. 10, pp. 21–46, 2018.

[7] D. K. Jain, R. Jain, X. Lan, Y. Upadhyay, and A. Thareja, "Driver distraction detection using capsule network," *Neural Computing & Applications*, vol. 33, no. 11, pp. 6183–6196, 2021.

[8] H. Al-Libawy, A. Al-Ataby, W. Al-Nuaimy, and M. A. Al-Taee, "Modular design of fatigue detection in naturalistic driving environments," *Accident Analysis & Prevention*, vol. 120, pp. 188–194, 2018.

[9] M. Moe Htay, "Feature extraction and classification methods of facial expression: a survey," *Computer Science and Information Technologies*, vol. 2, no. 1, pp. 26–32, 2021.

[10] C. S. Silveira, J. S. Cardoso, A. L. Lourenço, and C. Ahlström, "Importance of subject-dependent classification and imbalanced distributions in driver sleepiness detection in realistic conditions," *IET Intelligent Transport Systems*, vol. 13, no. 2, pp. 347–355, 2019.

[11] H. Zeng, C. Yang, H. Zhang et al., "A LightGBM-based EEG analysis method for driver mental states classification," *Computational Intelligence and Neuroscience*, vol. 2019, no. 4, pp. 1–11, 2019.

[12] P. Shi, M. Yu, X. Lu, and Z. Xuan, "Endurance brake classification control based on the generalized growth and pruning radial basis function neural network," *International Journal of Vehicle Structures & Systems*, vol. 9, no. 5, pp. 330–335, 2017.

[13] Z. Zhao, N. Zhou, L. Zhang, H. Yan, Yi Xu, and Z. Zhang, "Driver fatigue detection based on convolutional neural networks using EM-CNN," *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 7251280, 11 pages, 2020.

[14] M. Zhu, J. Chen, H. Li, F. Liang, L. Han, and Z. Zhang, "Vehicle driver drowsiness detection method using wearable EEG based on convolution neural network," *Neural Computing & Applications*, vol. 33, no. 20, Article ID 13965, 2021.

[15] V. Biradar, "Hardware implementation of driver fatigue detection system," *International Journal of Engineering and Technology*, vol. 9, no. 1, pp. 24–32, 2017.

[16] S. K. Satapathy, S. Dehuri, and A. K. Jagadev, "ABC optimized RBF network for classification of EEG signal for epileptic seizure identification," *Egyptian Informatics Journal*, vol. 18, no. 1, pp. 55–66, 2017.

[17] N. Jaworska, H. Wang, D. M. Smith, P. Blier, V. Knott, and A. B. Protzner, "Pre-treatment EEG signal variability is associated with treatment success in depression," *NeuroImage: Clinica*, vol. 17, no. C, pp. 368–377, 2018.

[18] A. K. Jaiswal and H. Banka, "Epileptic seizure detection in EEG signal with GModPCA and support vector machine," *Bio-Medical Materials and Engineering*, vol. 28, no. 2, pp. 141–157, 2017.

[19] V. Laprevote, L. Bon, J. Krieg et al., "Association between increased EEG signal complexity and cannabis dependence," *European Neuropsychopharmacology*, vol. 27, no. 12, pp. 1216–1222, 2017.

[20] W. Wang, S. Sun, L. Chao, and T. Zhichuan, "Recognition of upper limb motion intention of EEG signal based on convolutional neural network," *Journal of Zhejiang University*, vol. 7, no. 51, pp. 1381–1389, 2017.

[21] M.-G. Kim, H. Ko, and S. B. Pan, "A study on user recognition using 2d ecg based on ensemble of deep convolutional neural networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 5, pp. 1859–1867, 2019.

[22] T.-Y. Kim, S.-H. Kim, and H. Ko, "Design and implementation of BCI-based intelligent upper limb rehabilitation robot system," *ACM Transactions on Internet Technology*, vol. 21, no. 3, 2021.

[23] J. Trobl, M. Pioreck, and V. Kraja, "Methods for automatic estimation of the number of clusters for K-means algorithm used on eeg signal: feasibility study," *Lékar a Technika*, vol. 47, no. 3, pp. 81–87, 2017.

Retraction

Retracted: A Social-aware and Mobile Computing-based E-Commerce Product Recommendation System

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Xu and J. Wang, "A Social-aware and Mobile Computing-based E-Commerce Product Recommendation System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9501246, 8 pages, 2022.

Research Article

A Social-aware and Mobile Computing-based E-Commerce Product Recommendation System

Qin Xu [iD][1] **and Jun Wang**[2]

[1]*Department of Economics and Management, Wuhu Institute of Technology, Wuhu 241003, China*
[2]*Department of Information Construction and Management, Wuhu Institute of Technology, Wuhu 241003, China*

Correspondence should be addressed to Qin Xu; xuqin@whit.edu.cn

E-commerce product recommendation system can help users to find their own products quickly from a large number of products. To address the shortcomings of the current e-commerce product recommendation system, such as low efficiency and large recommendation errors, we designed an intelligent recommendation system based on social awareness and mobile computing. The behavioral characteristics of the current e-commerce product recommendation system are analyzed; the e-commerce product recommendation system is built according to the data processing technology of mobile computing, and the key technologies of the e-commerce product recommendation system are designed. The test results show that the proposed system overcomes the shortcomings of the traditional e-commerce product recommendation system, speeds up the speed of users to find the products they really need from a large number of products, improves the accuracy of e-commerce product recommendations, and the error of e-commerce product recommendations is much lower than that of the traditional, which has higher practical application value.

1. Introduction

With the continuous development of information technology, logistics technology, network technology, and intelligent technology, the number and variety of goods on the Internet has increased dramatically. Goods have been stored in huge quantities, and at the same time, people are trading goods on the Internet more and more frequently, and many types of e-commerce management systems have emerged [1–3]. Among the e-commerce management systems, the e-commerce product recommendation system is one of the most important subsystems, directly affecting whether users can efficiently search for the e-commerce products they want. Therefore, the design and research of recommendation systems have been an important research direction in the field of e-commerce applications [4–6]. The traditional e-commerce product recommendation system uses a single computer to manage all user requests and product data, but with the increasing number of users and product data, the shortcomings of the single-computer model have become

increasingly evident, mainly in the slow speed of the e-commerce product recommendation system, which makes it difficult to find the products that users really need within a short period of time [7–9].

The core task is to provide users with personalized product recommendation services by exploring the connection between users and products. Ultimately, achieving a win-win situation for both users and system owners [10, 11]. Large-scale e-commerce recommendation systems in the context of modern Internet applications are facing challenges in the following aspects: The data processing capacity of recommendation systems with a centralized architecture is limited. Stand-alone recommendation algorithms suffer from data processing scale limitations and processing efficiency problems [12]. In addition to the large number of users and products in large e-commerce systems, there are also many natural attributes of users and products, making it difficult to build accurate and effective models for high-dimensional users and products [13]. Business application requirements are often complex and variable, and user

concerns vary for different application scenarios, so recommendation systems based on fixed models and parameters often lack flexibility [14]. Recommendation models are often strongly correlated with data characteristics and application scenarios, which dictates that recommendation systems must integrate a variety of complementary recommendation technologies. Algorithms and models of existing solutions are relatively homogeneous, making it difficult to meet the mainstream needs of mainstream users while taking into account their individual needs [15].

In order to overcome the shortcomings of the stand-alone e-commerce product recommendation system, some scholars have designed a distributed processing technology-based e-commerce product recommendation system. The distributed processing system relies on the Internet to unify the management of multiple stand-alone machines, which significantly improves the efficiency of the e-commerce product recommendation system compared to the stand-alone working mode, but there are some shortcomings in the distributed e-commerce product recommendation system, such as the recommendation efficiency is still difficult to meet the user's requirements for a large quantity, and the e-commerce product recommendation error rate is high [5]. Mobile computing is a fast processing system based on a distributed system, with the advantages of parallelism, distribution, and robustness, and has been successfully applied in many big data processing fields [16].

2. Related Technical Analysis

2.1. Comparison of the Main Recommended Technologies.
The advantages and disadvantages of commonly used recommendation techniques, such as collaborative filtering, association rule mining, and knowledge-based empirical methods, are compared in Table 1.

2.2. Distributed Computing and Storage Technologies.
Google's papers on distributed infrastructure have had a huge impact on the industry, with ideas such as MapReduce [5] and GFS [17] providing key references for distributed computing and storage, of which Hadoop is an open-source implementation [18]. Hadoop avoids time-consuming data transfer problems when dealing with large-scale data by using mechanisms for distributed data storage and migrating code rather than migrating data; it allows the system to recover from node failures by using mechanisms for moderate data redundancy.

3. Dynamic Community Segmentation Based on Mobile Behavioral Similarity

Definition 1. : Mobile behavioral characteristics. Mobile behavior characteristics reflect the spatial and temporal distribution of users' movements in different spatial locations within a given time interval [19]. In this paper, we choose the spatial frequency and length of stay of users as the portrayal of mobile behavior characteristics, which can be expressed as

$$o\left(w_j, l_i\right) = \frac{f\left(w_j, l_i\right)}{f\left(w_j, l\right)} \times \frac{d\left(w_j, l_i\right)}{d\left(w_j, l\right)}, \tag{1}$$

where $o\left(w_j, l_i\right)$ represents the mobile behavioral characteristics of user w_j at a spatial location l_i, and $f\left(w_j, l\right)$ represents the frequency of visits by user w at f and at all spatial locations f, respectively, and $d\left(w_j, l_i\right)$ and $d\left(w_j, l\right)$ represent the length of stay of W, at l_i, and at all spatial locations z, respectively. The distribution of the characteristics of the movement behavior of user W, at all spatial locations, constitutes a q-dimensional vector (discrete spatial location-scale q) as

$$\mathbf{V}\left(w_j, l\right) = \left(o\left(w_j, l_1\right), o\left(w_j, l_2\right), \ldots, o\left(w_j, l_q\right)\right). \tag{2}$$

The distribution of mobile behavioral features used in mobile user community segmentation is a relative value calculation. In other words, the characteristic distribution portrays the user's "preference" for different spatial locations. In practice, the absolute values of the spatio-temporal distribution of different users in different spatial locations are different, i.e. $f\left(w_j, l_i\right)$ and $d\left(w_j, l_i\right)$ are significantly different, which is the basis for the concept of mobile activity of users within a community.

Definition 2. : Mobile activity. Mobile activity is relative to the distribution of mobile behavioral characteristics, which are essentially relative values, while mobile activity is an absolute value [20]. Mobile activity is defined as the product of the cumulative frequency of visits and length of stay of a user W, at different spatial locations, expressed as

$$a\left(w_j, l\right) = \sum_{i=1}^{q} f\left(w_j, l_i\right) \times \sum_{i=1}^{q} d\left(w_j, l_i\right). \tag{3}$$

Definition 3. : Mobile spatio-temporal communities. Given a discrete time interval c_i and a dataset of historical movement trajectories of a set of users w, the set of users is divided into a finite number of communities based on the calculation of the characteristics of the movement behavior of different users at different time intervals c_i and different spatial locations is expressed as follows:

$$P = \{p_1, p_1, p_2, \ldots, p_k\}, \tag{4}$$

where $p_i = (w_{i1}, w_{i2}, \ldots, w_{ix})$, w_{ij} corresponds to the ith segmented community, and users within the same community have similar mobile behavioral profile distribution, while users within different communities have different mobile behavioral profile distribution [21]. The following section describes the process of dividing mobile communities in detail. Using the mobile behavioral profile formula, the preference profile of each participating user for different spatial locations can be derived $o\left(w_j, l_i\right)$, for each user w, the profile is calculated at different spatial locations, and the resulting preference profile is expressed in the form of a matrix as follows:

TABLE 1: Comparison of the main recommended technologies.

	Advantage	Shortcoming
Collaborative filtering technology	Discover new and different interests and do not depend on domain knowledge. With the passage of time and the accumulation of data, the effect is getting better and better. The recommendation process has a high degree of personalization and automation and can deal with complex unstructured objects	There are typical problems such as scalability, sparsity and cold start, and the recommendation quality depends on the historical data set
Association rule mining	Can discover new and different points of interest, independent of domain knowledge	Rule extraction is difficult and time-consuming, and the degree of personalization is low
Knowledge experience-based approach	It can consider nonproduct attributes, reflect user needs, and make up for the lack of user knowledge and experience	Knowledge and experience are difficult to obtain, and recommendation is static

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1q} \\ m_{21} & m_{22} & \cdots & m_{2q} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nq} \end{bmatrix}, \tag{5}$$

where $m_{j,i} = o(w_j, l_i)$, the cosine similarity is calculated for any 2-row vector of matrix M by the following formula:

$$\cos \theta_{j1,j2} = \frac{\sum_{i=1}^{n} \left(o(w_{j1}, l_i) \times o(w_{j2}, l_i) \right)}{\sqrt{\sum_{i=1}^{n} \left(o(w_{j1}, l_i)^2 \right)} \times \sqrt{\sum_{i=1}^{n} \left(o(w_{j2}, l_i)^2 \right)}}, \tag{6}$$

The similarity between users w_{j1} and w_{j2} in all spatial locations based on the distribution of their mobile behavioral characteristics is represented by $\cos \theta_{j1,j2}$. Based on this similarity, the k-means clustering algorithm is used to classify the n participating users into communities, resulting in seven different mobile communities. As mentioned above, the mobile activity calculation is used to identify a unique community organizer in each of the divided mobile communities m as the end-user super-user for task reception and data upload with the MCS platform, and as the task coordinator between different user subordinates within the community [22]. In order to indirectly group the spatial locations by communities, we use the average value of the spatial preference vector of all users in the community to represent the community's preference for spatial locations. If the community structure is $p_k = \{w_{i1}, w_{i2}, \ldots, w_{ip}\}$, then the community preference for spatial location $\mu(p_k, l)$ is expressed as

$$\mu(p_k, l) = \frac{1}{p} \sum_{x=i1}^{ip} V(w_x, l)$$

$$= \frac{1}{p} \sum_{x=i1}^{ip} [o(w_x, l_1), \ldots, o(w_x, l_n)]. \tag{7}$$

4. Merchandise Recommendations for Mobile Computing

The traditional e-commerce product recommendation system uses a single-computer model, where the user's e-commerce product tasks and related data are all stored on a single computer, and all work is completed on this computer, making the intelligent recommendation of e-commerce products quite time-consuming and unable to meet the current requirements of e-commerce product development [23]. Mobile computing introduces parallelism and task decomposition techniques on the basis of distributed processing systems, where a large-scale task is processed in pieces, resulting in many subtasks, each of which is carried out by one computer (node), so that the processing results of each subtask are obtained, and finally, the results of the subtasks are fused to obtain the final processing results, shown in Figure 1.

In an e-commerce product intelligent recommendation system, the recommendation algorithm is very critical, which directly affects the effect of e-commerce product intelligent recommendation. Therefore, this paper presents the specific design of the e-commerce product intelligent recommendation algorithm.

Let there be N users, which are denoted as User $= \{u_1, u_2, \ldots, u_j, \ldots, u_N\}$, and all e-commerce products are denoted as Item $= \{i_1, i_2, \ldots, i_j, \ldots, i_p\}$, and p denotes the total number of e-commerce products; there are M user tags, which can be denoted as Tag $= \{t_1, t_2, \ldots, t_j, \ldots, t_M\}$ so that the frequency of users clicking on e-commerce products and the corresponding tag values of e-commerce products describe the user characteristics, which are calculated as follows [24]:

$$V_{ut_l} = \mathrm{tf}_{\text{user rax}} \cdot \mathrm{idf}_{\text{user}_{\text{tus}}} \cdot \left\{ \frac{n_{ut_l}}{n_{ut}} \log\left(\frac{N}{N_{ul}}\right) \cdots \frac{n_{ut_i}}{n_{ut}} \log\left(\frac{N}{N_n}\right) \cdots \frac{n_{ul_N}}{n_{ut}} \log\left(\frac{N}{N_{ut_N}}\right) \right\}. \tag{8}$$
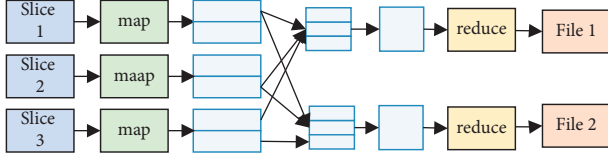
Figure 1: Working principle of e-commerce product recommendation system for mobile computing.

Let n_{ut_i} and n_{ut} be the number of times a user uses t_i and the total number of tags used by a user respectively, then the frequency of tags and the popularity of tags to users are calculated as follows:

$$\text{tf}_{\text{user}_l ag} = \frac{n_{ut_i}}{n_{ut}},$$

$$\text{idf}_{\text{user}_l ag} = \log\left(\frac{n}{n_{ut_N}}\right). \tag{9}$$

The feature vector describes the importance of e-commerce good and is calculated using the following metric:

$$V_{ut_j} = \text{tf}_{\text{item}_{ux}} \cdot \text{idf}_{\text{item}_{ta}} \cdot \left\{ \frac{n_{it_l}}{n_{it}} \log\left(\frac{P}{N_{it_i}}\right) \cdots \frac{n_{i_i}}{n_{it}} \log\left(\frac{P}{N_{i_j}}\right) \cdots \frac{n_{it_H}}{n_{it}} \log\left(\frac{P}{N_{it_N}}\right) \right\}. \tag{10}$$

where $\text{tf}_{\text{item}_{ux}}$ and $\text{idf}_{\text{item}_{ta}}$ denote the frequency and importance of the use of the e-commerce good, respectively; n_{it_l} is the number of times e-commerce good i is used for the jth time; n_{it} denotes the total number of uses of e-commerce good i.

The user preference matrix for e-commerce goods is calculated as follows:

$$\mathbf{V}_{uji_k} = \sum_{i=1}^{M} \mathbf{V}_{uj} \times \mathbf{V}_{i_k t} \tag{11}$$

In which $u_j \in U, j = 1, 2, \ldots, N; i_k \in I, k = 1, 2, \ldots, P$.

The e-commerce product preference characteristics of users u_j are

$$V_{u_j} = \left(V_{uj,i}, V_{uji}, \ldots, V_{uji_k}, \ldots, V_{u,i_i}\right), \tag{12}$$

where V_{uji_k} is the preference level of u_j for e-commerce products. Based on the user's preference vector for e-commerce products, the user e-commerce product preference matrix is established as

$$\mathbf{V}_{N \times P} = \begin{bmatrix} V_{u_1 i_1} & \cdots & V_{u_1 i_k} & \cdots & V_{u_1 i_p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ V_{u_j i_1} & \cdots & V_{u_j i_k} & \cdots & V_{u_j i_p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ V_{u_N i_1} & \cdots & V_{u_N i_k} & \cdots & V_{u_N i_p} \end{bmatrix}. \tag{13}$$

Similarity describes the degree of similarity between two e-commerce products, and intelligent recommendation algorithms obtain the similarity of e-commerce products based on cosine similarity [25]. The feature vector of an e-commerce product can be expressed as

$$\mathbf{I}_k = \left(n_{k1}, n_{k2}, \ldots, n_{ki}, \ldots, n_{kL}\right), \tag{14}$$

where n_{ki} is the value of the t_i-tagged e-commerce item i_k normalized. The full e-commerce product information can be described using the e-commerce product-specific vector matrix $\mathbf{I}_{k \times k}$ as

$$\mathbf{I}_{k \times k} = \begin{bmatrix} n_{11} & \cdots & n_{1k} \\ \vdots & \ddots & \vdots \\ n_{k1} & \cdots & n_{kk} \end{bmatrix}. \tag{15}$$

Cosine similarity between e-commerce products based on eigenvectors is expressed as follows:

$$\text{sim}\left(i_j, i_k\right) = \cos\left(\mathbf{I}_j, \mathbf{I}_k\right) = \frac{\mathbf{I}_j \cdot \mathbf{I}_k}{\left|\mathbf{I}_j\right| \times \left|\mathbf{I}_k\right|}. \tag{16}$$

(1) Let us divide an e-commerce product smart recommendation task and map each subtask to a corresponding mobile computation point via the map program [26]. (2) At each node, the user's preference matrix for the e-commerce product is calculated. (3) At each node, the similarity of e-commerce products is calculated and an e-commerce product similarity matrix is built. (4) Based on the user's historical search data for e-commerce products, each node searches the historical e-commerce products that user u has clicked on, and calculates the preference value between the user and the e-commerce product. (5) At each node, the preference values between users and e-commerce products are ranked, and the top N e-commerce products are selected as the intelligent recommendation results and the recommendation results of each node are obtained. (6) The recommendation results of each node are output to reduce, and the best N e-commerce products are finally recommended to the user through reduce fusion [27, 28].

The intelligent recommendation process of e-commerce products is shown in Figure 2.

5. Experimental Analysis and Testing

5.1. E-Commerce Product Datasets. To test the performance of the mobile computing-based e-commerce product recommendation system, an e-commerce management system with 50,000 e-commerce products and 2,000 users was selected as the test object. 80% of e-commerce product and user data were randomly selected as the training set, while the other data were used as the validation set. The performance of the e-commerce product intelligent recommendation system
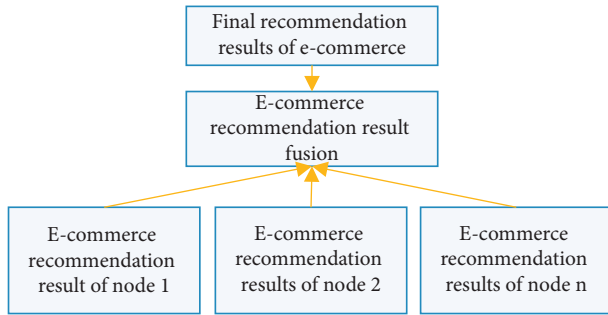
Figure 2: Intelligent recommendation process for e-commerce products.

is evaluated by the accuracy and completeness rates, where the accuracy rate represents the accuracy of e-commerce product intelligent recommendation and the completeness rate represents the reliability of e-commerce product intelligent recommendation.

5.2. Comparison of Recommended Results. The accuracy and completeness rates of the e-commerce product intelligence recommendations are shown in Figures 3 and 4, respectively. It is clear from Figures 3 and 4 that the accuracy and completeness rates of the mobile computing-based e-commerce product intelligent recommendation are higher than those of the traditional e-commerce product intelligent recommendation. This is mainly because the e-commerce product intelligent recommendation algorithm in this paper takes into account users' preferences for e-commerce products, which improves the accuracy of the e-commerce product intelligent recommendation, and the reliability and stability of the e-commerce product recommendation are higher.

The recommendation time of the e-commerce product recommendation system is shown in Figure 5. As can be seen from Figure 5, the intelligent recommendation time for e-commerce products based on mobile computing has been significantly reduced, overcoming the shortcomings of the traditional e-commerce product intelligent recommendation system in terms of long recommendation times and low work efficiency, and helping users to find the e-commerce products they really need more quickly.

5.3. Social Perception Effect. This paper uses the WTD public dataset, which records approximately 11 weeks of communication data between 275 users with PDAs (personal digital assistants) and deployed APs on the UC San Diego campus, as experimental data. This dataset was used to validate the performance of the proposed group intelligence-aware social task distribution approach for mobile social networks. Due to the sparsity of different user data in the WTD dataset, 140 AP nodes are selected as spatial locations and 68 users are selected as participating users in this experiment. Distribution of mobile community behavioral features and the user connectivity network based on closeness is first tested.
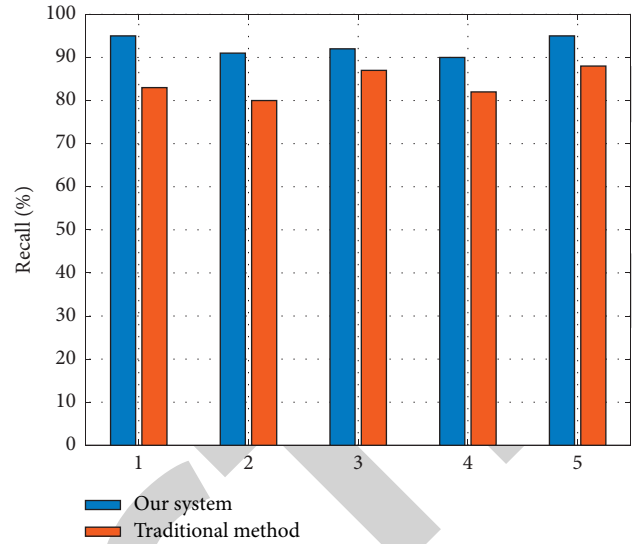
Figure 3: Recall of e-commerce product intelligence recommendation system.
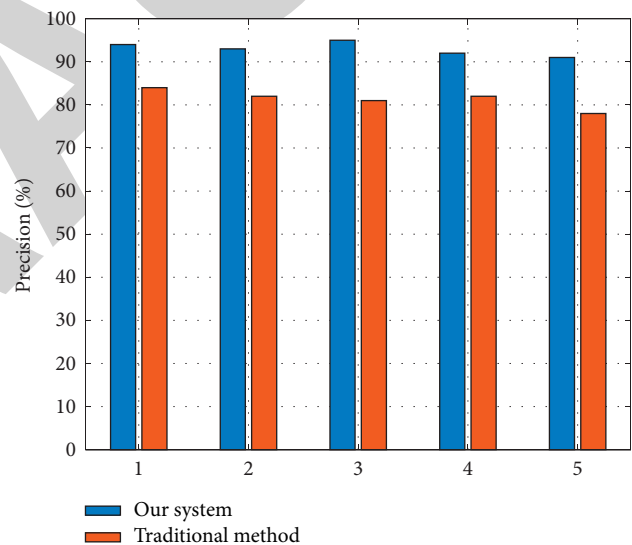
Figure 4: Completeness rate of intelligent recommendation systems for e-commerce products.

The task execution effect, the number of task distributions, and the energy consumption of task recovery were tested by randomly generating 100–400 mobile group intelligence-aware tasks, with the spatial location of the tasks randomly selected from 140 AP points and the time interval randomly selected from 0 : 00 to 24 : 00. The relevant parameters are set as follows: the user fails to perform the task with a probability of 0.3, the community organizer respects a new performer with a probability of 0.5, and the task is passed through the social network with a probability of 0.5; in the case of passing the task, the person to whom the task is passed performs or refuses to perform the task with a probability value of their closeness. The energy consumption coefficient for data transfer via direct networks is set to 1, and the energy consumption for data transfer via
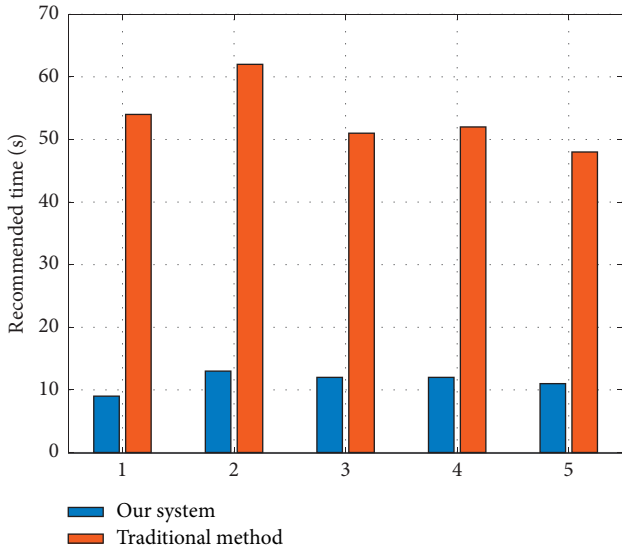
Figure 5: Recommendation time of an intelligent recommendation system for e-commerce products.
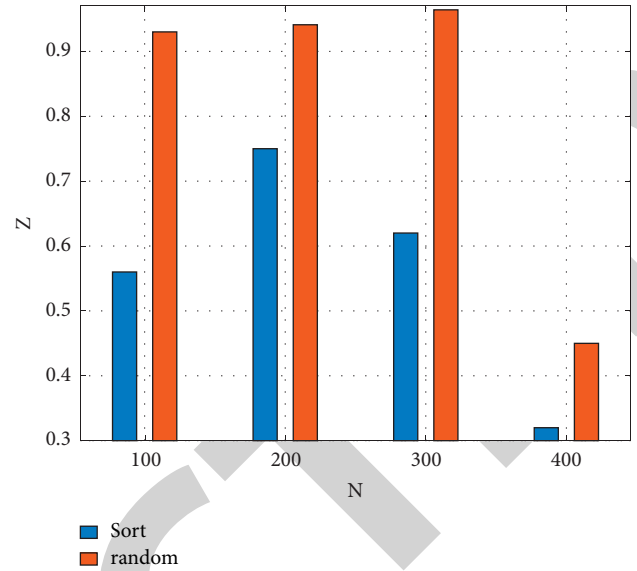
Figure 7: MCS task distribution experiment I.
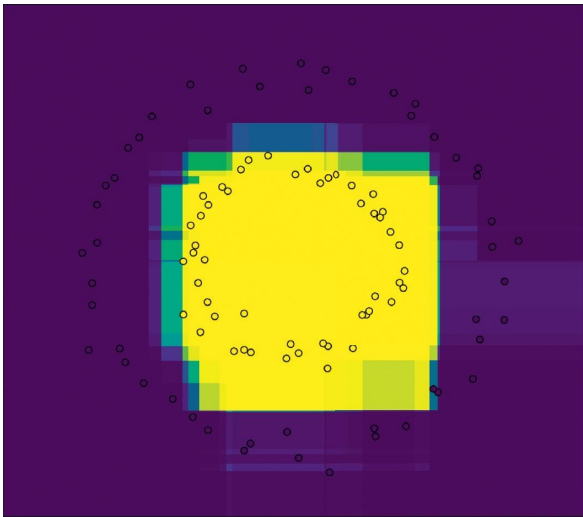
Figure 6: Social connection network based on intimacy.

Figure 8: MCS task distribution experiment II.

opportunistic networks is set to 0.4. The specific distribution of the socially connected network in terms of closeness is shown in Figure 6.

Figure 7 shows a comparison experiment between the random distribution strategy and the traditional sequential distribution strategy, as well as an experiment with different intervention strategies after task execution failure, where vs is the number of group wisdom-aware tasks and z is the task completion rate. In the comparison between the random distribution strategy and the traditional sequential distribution strategy, there is no replacement of users for the failed task, and if the selected user does not appear in the specified spatial area on the second day at the corresponding time interval, the task fails. It can be seen from the figure that the proposed random distribution strategy significantly outperforms the traditional sequential distribution strategy.

Figure 8 shows the distribution statistics of the tasks under the task execution failure intervention experiment and the results of the perceptual data recovery energy simulation experiment, where B is the number of task distributions, u is the energy consumption statistics, and IV is the number of tasks. The statistics of the number of task distributions under the failed intervention experiment show that the number of distributions is lowest for the community-based distribution and the social intimacy-based user delivery, followed by the community-based distribution and the replacement within the failed user community, and the traditional "platform one user" full user reselection. The reason for this is that user delivery and distribution based on social intimacy has a high rate of task acceptance and

execution. In addition to the above experiment, a questionnaire was conducted to simulate a real-life campus environment. This figure is generally consistent with the results of the WTD test. The reason for this is that the dynamic community segmentation based on spatio-temporal mobile behavior clusters users with similar mobile behavior so that suitable users can be searched for within a relatively small range of available users for replacement. The results of the perceived data recovery energy consumption simulation show that the optimal replacement strategy within the failed user community is, in order, social relationship-based intimate user delivery and distribution and the traditional "platform one user" full user reselection. The reason for this is that the optimal intracommunity replacement strategy maximizes the use of opportunistic mobile social networks for short-range communication and therefore consumes the least amount of energy, while the social relationship-based intimate user transfer and distribution comes second and the traditional "platform one user" full user reelection uses 3 G/4 G communication for data upload/download, which consumes the most energy. The energy consumption is the highest.

6. Conclusions

In order to overcome the shortcomings of the current e-commerce product recommendation system, this paper proposes a mobile computing-based e-commerce product recommendation system. Firstly, to address the shortcomings of the current e-commerce product recommendation system based on the stand-alone working mode, such as low efficiency and slow speed, the mobile computing working mode was introduced to decompose the e-commerce product recommendation task and carry out distributed and parallel processing; secondly, to address the problem of large recommendation errors in the e-commerce product recommendation system, the e-commerce product intelligent recommendation system integrates the user's preference for products. The results show that this paper's e-commerce product intelligent recommendation system works efficiently and has high accuracy in recommending e-commerce products, which can provide valuable reference advice for users when conducting e-commerce product transactions.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] W. S. Yang, H. C. Cheng, and J. B. Dia, "A location-aware recommender system for mobile shopping environments," *Expert Systems with Applications*, vol. 34, no. 1, pp. 437–445, 2008.

[2] H. Liu and Z. Liang, "Implicit rating model in M-commerce recommendation system," in *Proceedings of the 2009 International Conference on Computational Intelligence and Software Engineering*, pp. 1–4, IEEE, Wuhan, China, December 2009.

[3] R. Colomo-Palacios, F. J. García-Peñalvo, V. Stantchev, and S. Misra, "Towards a Social and Context-Aware mobile Recommendation System for tourism," *Pervasive & Mobile Computing*, vol. 38, pp. 505–515, 2016.

[4] N. Manouselis and K. Verbert, "Layered evaluation of multicriteria collaborative filtering for scientific paper recommendation," *Procedia Computer Science*, vol. 18, no. 1, pp. 1189–1197, 2013.

[5] I. O. Pappas, P. E. Kourouthanassis, M. N. Giannakos, and V. Chrissikopoulos, "Shiny happy people buying: the role of emotions on personalized e-shopping," *Electronic Markets*, vol. 24, no. 3, pp. 193–206, 2014.

[6] R. Lacuesta, G. Palacios-Navarro, C. Cetina, L. Peñalver, and J. Lloret, "Internet of things: where to be is to trust," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–16, 2012.

[7] E. D. Cavanagh, "The rule of reason Re-examined," *The Business Lawyer*, vol. 67, no. 2, pp. 435–469, 2012.

[8] Q. Lu and F. Guo, "A novel e-commerce customer continuous purchase recommendation model research based on colony clustering," *International Journal of Wireless and Mobile Computing*, vol. 11, no. 4, pp. 309–317, 2016.

[9] F. Long, "Improved personalized recommendation algorithm based on context-aware in mobile computing environment," *Wireless Communications and Mobile Computing*, vol. 2020, no. 1, 10 pages, Article ID 88, 2020.

[10] S. Rotchanakitumnuai and M. Speece, "Corporate customer perspectives on business value of Thai internet banking," *Social ence Electronic Publishing*, vol. 5, pp. 270–286, 2004.

[11] Y. M. Wang, R. L. Wang, and D. X. Xu, "Mobile E-commerce personalized information recommendation model," *Applied Mechanics and Materials*, pp. 2136–2139, 2014.

[12] J. Lin, Y. Lu, B. Wang, and K. K. Wei, "The role of interchannel trust transfer in establishing mobile commerce trust," *Electronic Commerce Research and Applications*, vol. 10, no. 1–6, pp. 615–625, 2011.

[13] R. Rianto, L. E. Nugroho, and P. I. Santosa, "Analysis of E-commerce user behavior of Indonesian students: a preliminary study of adaptive E-commerce," *Studies in Computational Intelligence*, vol. 595, pp. 365–375, 2015.

[14] K. Meehan, T. Lunney, K. Curran, and A. McCaughey, "Context-aware Intelligent Recommendation System for tourism," in *Proceedings of the IEEE International Conference on Pervasive Computing & Communications Workshops*, pp. 328–331, IEEE, San Diego, CA, USA, March 2013.

[15] M. Eddahibi, A. Nejeoui, and C. E. Cherkaoui, "Towards an intelligent and deeply automated shopping bot," *International Journal of Computer Application*, vol. 45, no. 24, pp. 21–27, 2012.

Research Article

Intelligent Light Control System Based on Zigbee

Xinxin Wang and Lijun Wang [ID]

School of Mechanical Engineering, North China University of Water Resources and Electric Power, Zhengzhou, China

Correspondence should be addressed to Lijun Wang; wanglijun@ncwu.edu.cn

This system uses Freescale i.MX283 chip as the main control core of the control panel and LPC824 as the main control core of the light control point. ZigBee communication local area network is built. The ZigBee network is used to search all the nodes and display them on the control interface stably. The controller can selectively control the node and communicate with ZigBee. It supports simple and convenient human-computer interaction interface. The control panel can control RGB lamps of lamp control nodes as follows: control the light on and off, adjust the brightness of the light, toggle the color of the light, control the gradient of the light, control the flashing of RGB lamp, control timing switch of RGB lamp, realize the scene effect of the light, etc. After testing, the system is stable and applied to many lighting control occasions.

1. Introduction

With the rapid development of science and technology, the development trend of wireless communication combined with intelligent control systems is more and more obvious, such as smart homes and the Internet of things. They all take the system as a carrier and rely on a way of communication to control the state and exchange data. The control side controls the state (foreground effect), and the node changes its state (background effect). The system mainly processes and calculates various complex data and realizes the algorithm. Communication is mainly used for data transmission and the exchange of feedback. In this way, it can not only realize highly complex algorithm functions but also support a human-computer interaction interface which is convenient for users. The intelligent lighting control system based on ZigBee communication integrates computer technology, mechanical control theory, sensor technology, system engineering, and artificial intelligence technology. It has a wide range of applications and is the development trend of Internet of things communication in the future [1].

2. Overall Design Scheme

This design selects i.MX283 as the control interface; its core is an ARM9 processor. The processing speed is fast enough because it has enough RAM space; it can also run common operating systems such as UCOS, LINUX, and AWorks. It has its own ZigBee minimum system module supporting the ZigBee protocol. The LPC824 of Freescale is selected as the control node, and the processor is arm CortexM0+ which is a low-power controller with a processing speed of about 30 m. For the light control node, this speed is large enough, and it has its own ZigBee minimum system module that supports ZigBee protocol and can realize various state controls of RGB [2].

The software design is mainly divided into control panel and light control node. The control panel provides a simple human-computer interaction control interface for users who can click the interface control icon to send out the corresponding control signal according to the needs of their own functions. Then, the data layer will send the control signal to the light control node for execution through ZigBee protocol communication. The light control node continuously

queries and receives the control signal sent by the control panel and then analyzes the command and executes the corresponding control command to achieve the corresponding functions.

The control panel mainly involves five thread tasks [3]:

(1) GUI tasks: used to create and display function icons. When the user clicks the function icon, the callback function will be called to complete the corresponding function. For example, select and switch the target control node, send the control status command, receive the temperature and humidity data, mailbox data, and real-time display, receive the surrounding node information and display node-related information, and modify the display name of the node to facilitate memory.

(2) Temperature and humidity acquisition task of DHT11: mainly collect the current temperature and humidity, and send the data to the GUI interface through the temperature and humidity data mailbox.

(3) Receive control signal task of light control node: receive the control information data from the GUI task layer in real time, and then, the task is sent by ZigBee.

(4) ZigBee search task: search the surrounding node information, and send all node information to GUI.

(5) ZigBee sends control status task: sends control information data to the lamp control node.

(6) Light control node task: the light control node is equivalent to a state machine. ZigBee protocol is used to receive the control information from the control interface in real time and then analyze and execute the control information.

3. Software Design

The software development of intelligent lighting control system is based on eclipse software as the development environment, C language as the main programming language, and Freescale chip ARM9 and cortex M0+ as the core processor. The software module needs to use the task allocation, task scheduling, communication mechanism between tasks, GUI, GPIO port control, UART module, PWM module, and ZigBee communication protocol of real-time system AWorks.

3.1. ZigBee Networking and Communication Process. The coordinator can create a unique WLAN PANID. Routers and terminals can join a wireless LAN PANID, that is, network access. The PANID created by the coordinator and the PANID added by the router and terminal are preset in the program and are not automatically allocated by the module. Therefore, in the process of network creation and network addition, it is necessary to specify which LAN to create or need to join, so as to ensure the effectiveness of data transmission.

The coordinator module establishes the wireless local area network. This PANID is unique. When selecting the channel, select the channel without a local area network nearby (there are 16 channels in 2.4 G bandwidth to choose from 11 to 26) which can prevent competition and increase the stability of the signal [4].

When the coordinator creates a LAN and a module (router or terminal) joins the LAN, the node can communicate with the router. Terminals cannot communicate directly with each other and need to pass through routers to act as an intermediary. Sending data must comply with ZigBee protocol and can be developed using ZigBee protocol stack provided by ZigBee manufacturer. The protocol stack is like a development library written according to the provisions of the ZigBee protocol. Developers only need to develop through the API functions provided by the library functions so as to realize the sending and receiving of wireless data [5].

3.2. Interface Layer Design of the Control Panel

3.2.1. GUI Control Interface. emWin and ucGUI are the same; GUI graphics library is like a function encapsulation library which provides developers with various interfaces for drawing API. Understand API functions and precautions to carry out preliminary development, and emWin interface application structure is clear; each page is composed of three parts: (1) control structure array, (2) callback function, and (3) page creation function. Then, there is the window and control management which are mainly composed of (1) control handle, (2) control ID, and (3) message [6, 7].

(1) Control Array. View all the controls used in the current interface. The main function is to create and initialize the information of related controls such as the abscissa and ordinate of the control displayed in the current interface, the size, and name of the control.

(2) Callback Function.

```
static void _dialog(massge *p_msg) {
switch (p_msg-> id) {
case init://Reset control properties (size, color, etc.)
case parent://Operations trigger message processing such as clicking the//button, clicking the edit box
case timer://Timer trigger operation, timer will execute the program here
case paint://Window redraw message, display text in the window or draw lines,//rectangles, circles
}
}
Page creation function:
void task_ main(void)
{
gui_ Init();//initialization interface
create_ Win();//create interface
```

while (1) {delay (20); }//call GUI_ Delay function delay 20 ms

}

Each page has its own callback function which contains all the functions that the control needs to implement. Any operation of the interface will trigger and call this callback function and then process it in the program.

3.2.2. GUI Interface Design. The intelligent light control system based on ZigBee mainly includes the display and selection of node, the switch of node light, the color of light, the brightness of light, the flicker of light, the scene effect of light, the timing of light, and the gradient of light. The control selection mainly includes button control, checkbox control, slider control, and text control [8, 9].

(1) Creation and Start of Interface.

```
void gui_task(void * p_arg)/* GUI task function
entry */
{
WM_HWIN Win;
GUI_Init();/* Initialization interface */
Win = CreateSmsrt_Control(WM_HBKWIN);/*
Create interface */
while(1) {
GUI_Delay(20);/* Waiting for task scheduling */
}
}
```

Button control is used for the selection of light color. When user presses button control, GUI task will call callback function to create a color selection interface.

The program framework of button control is as follows:

```
static void dialog(massge * p_msg) {/* GUI task
callback function */
switch (p_msg ->msg_id) {/* Type judgment of
callback */
case notify_parent:/* The callback type is operation
trigger message */
id = get_id(p_msg ->win_src);/* Get interface con-
trol type */
ncode = p_msg -> data.v;/* Operational data */
switch(id) {/* Control type selection */
case botton_0:/* botton_0 Control type */
switch(ncode) {
case click:
break;
case reaessed;/* Click operation */
create_colo_win();/* Create a color selection interface
*/
break;
}
break;
```

}}}

(2) Checkbox Control. Switch for light, scene effect, timing function, and gradient function. It is equivalent to a check box. Open a function (such as scene effect function); click the scene effect option box to check it. Then, the GUI task will call the callback function to execute the selected condition options of the checkbox control, and the LED status data mailbox will be sent out to control the lighting effect of the light control node [10, 11].

The program framework of the checkbox control is as follows:

```
static void dialog(massge * p_msg) {/* GUI task
callback function */
switch (p_msg ->msg_id) {/* Type judgment of
callback */
case notify_parent:/* The callback type is operation
trigger message */
id = get_id(p_msg->win_src);/* Get interface control
type */
ncode = p_msg -> data.v;/* Operational data */
switch(Id) {/* Control type selection */
case check_1:/* check_ 1 control type */
switch (ncode) {
case click:
break;
case reaessed:{/* With ID_ CHECKBOX_ 1 click
operation */
send_mailbox(mail_type,&data, wait_time);/* Send
various data messages */
break;
}
break;
}} }
```

(3) Slider Control. Used for brightness adjustment of light and flicker frequency adjustment of light. It is like people seeing progress bars, but it has its own length range and size. It is like a ruler; drag it randomly, and the current dimension value will be represented at the last stop. Then, the size returned by the ruler can be used to adjust the intensity and flicker frequency of the RGB lamp which is proportional to the length of the size [12, 13].

The framework of the glider control is as follows:

```
static void dialog(massge * p_msg) {/* GUI task
callback function */
switch (p_msg ->msg_id) {/* Type judgment of
callback */
case notify_parent:/* The callback type is operation
trigger message */
id = get_id(p_msg->win_src);/* Get interface control
type */
```

ncode = p_msg - > data.v;/ ∗ Operational data ∗ /

switch(id) {/ ∗ Control type selection ∗ /

case slider_1:/ ∗ SLIDER _ 1 control type ∗ /

switch (ncode) {

case clicked:

break;

case released:/ ∗ SLIDER _ 1 click operation ∗ /

hItem = get_dialog_item(p_msg- > win, slider_1); / ∗ Get handle ∗ /

slider_set_width (hitem,15);/ ∗ Sets the size of the slider ∗ /

slider_set_range(hitem,0,9);/ ∗ Set the scale range of slider ∗ /

val = slider_get_value(hitem);/ ∗ Gets the current size value ∗ /

send_mailbox(mail_type,&data, wait_time);/ ∗ Send data message ∗ /

break;

}

break;

}}}

(4) Timer Mechanism of GUI. GUI handles corresponding events through the callback function. When time is up, it will execute an option callback function of the timer, and the timer will be set once. When the timer is used again, the timer count must be enabled again. This design mainly has two timers: one is responsible for the timing, receiving temperature and humidity data information every 2 seconds, and displaying it on the interface. The other is responsible for receiving the number of surrounding nodes every 3 seconds, dynamically increasing and deleting the number of nearby nodes, and displaying them on the interface [14].

Timer program framework is as follows:

static void dialog(massge ∗ p_msg) {/ ∗ GUI task callback function ∗ /

switch (p_msg - > msg_id) {/ ∗ Type judgment of callback ∗ /

case timer:/ ∗ The callback type is timer ∗ /

if (get_timer_id(p_msg - > data.v) = = 0) {/ ∗ TIMER0 ∗ /

receive_mailbox(mail_type, p_data,wait_time);/ ∗ Receive temperature and humidity data mailbox ∗ /

restart_timer(p_msg - > data.v, 2000);/ ∗ Restart the timing for 2 seconds ∗ /

}

if (get_timer_id(p_msg - > data.v) = = 1) {/ ∗ TIMER 1 ∗ /

receive_mailbox(mail_type, p_data,wait_time);/ ∗ Receive data mailbox of surrounding nodes ∗ /

restart_timer(p_msg - > data.v, 3000);/ ∗ Restart the timing for 3 seconds ∗ /

} }}

3.3. Design of Data Layer of the Control Panel.

This system needs frequent data transmission and data transmission between tasks. So the data processing layer is designed. All data will be collected, sent, or accepted directly or indirectly through the data layer. The main data involved are temperature and humidity data, status command data of light control node, number of surrounding nodes, control information data sent by ZigBee, search node queue data, current node information list, and modified name node information list. The main tasks involved are as follows: ZigBee searches the information of nearby nodes, sends the control data to the light control node, receives the light control node data from the interface layer, and collects the temperature and humidity in real time and sends it to the interface layer and interface touch screen [15, 16]. The overall structure of the control panel is shown in Figure 1.

3.3.1. Realization of Touch Screen Tasks.

It mainly completes the initialization and calibration of LCD and monitors whether there are touch screen operations in real time. When there are touch screen operations, it will call the interface callback function to perform the corresponding functions. Its task level of 2 is the highest so as to successfully monitor all touch screen operations [17]. The flow chart of touch screen detection is shown in Figure 2.

3.3.2. Realization of Temperature and Humidity Real-Time Acquisition.

According to the protocol specification of DHT11, read the measured temperature and humidity and then send the data to the interface layer display by e-mail through intertask communication. Mission priority is 7. The driving steps of the DHT11 temperature and humidity sensor are as follows:

(1) After the DHT11 module is powered on, wait for a second before sending the instruction to the module and set the data pin of DHT11 to the input state to detect the external signal.

(2) The controller sends out a start signal (low level) and then waits for the temperature and humidity module to respond to the controller [18].

(3) After the temperature and humidity module generates the response signal, the high level of 80 microseconds is output to tell the controller to prepare the data to receive.

(4) After the preparatory work is completed, DHT11 starts to send data to the controller; the controller only needs to receive it. The data output by DHT11 has 40 bits including 16 bits of temperature data, 16 bits of humidity data, and 8 bits of calibration data.
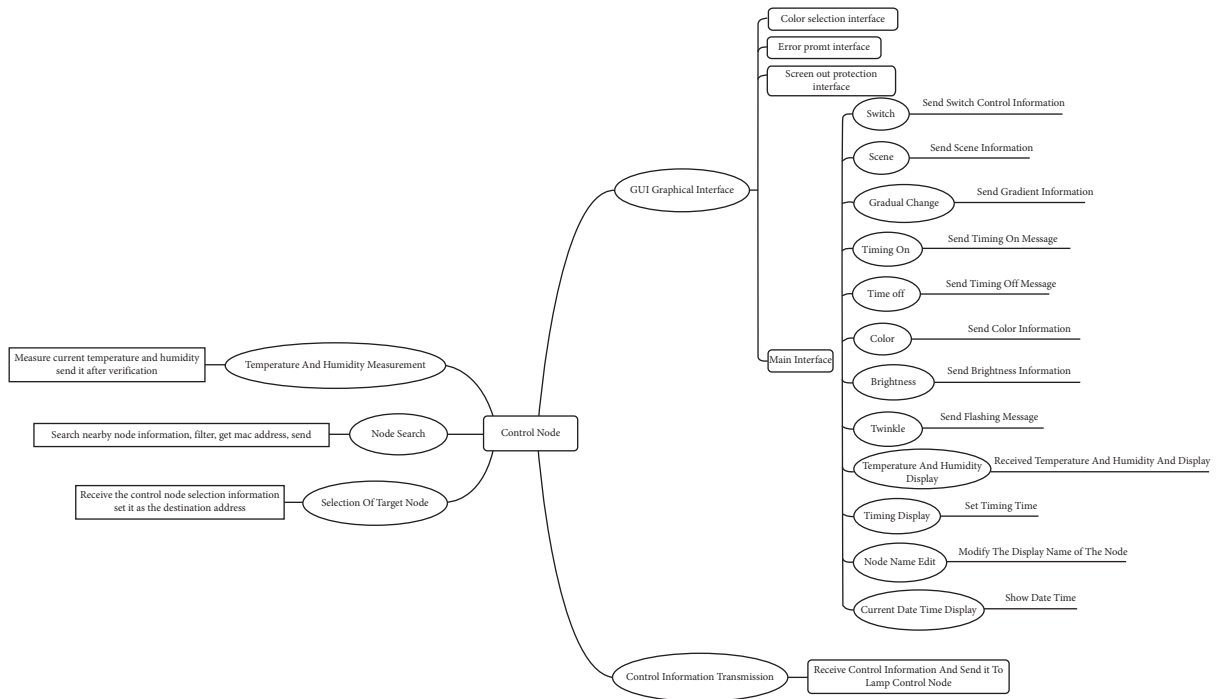
Figure 1: General structure of the control panel.
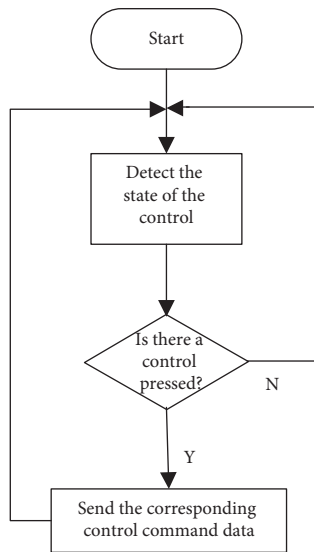
Figure 2: Flow chart of touch screen detection.

(1) Data Format. Because DHT11 adopts single bus communication, it can only transmit data in one direction at the same time. It transmits bit by bit (generally used to transmit digital data). Because data are transmitted through high and low levels, it is necessary to specify the form of data. The format of data "0" is 50 ms low level + 28 ms high level and the format of data "1" is 50 ms low level + 70 ms high level.

(2) Controller Parsing Data. The data received by the controller is continuous level, so developers need to convert the level into effective data. According to the data format, there are many kinds of parsing methods. The method used in this design is as follows: first, wait for the low level to pass, that is, wait for the data line to pull high, then delay for 60 us, because 60 us is greater than 28 us and less than 70 us, and then detect whether the data line is high at this time. If it is high, the data is determined as 1; otherwise, it is 0 [19, 20].

(3) Precautions. There will be a lot of delay waiting in the process of starting the temperature and humidity sensor and reading the data. Active delay in the operating system causes the operating system to perform other tasks with high priority. If the temperature and humidity measurement task is delayed, but the operation system does not return (because there are high priority tasks to be processed), the temperature and humidity data read will be incorrect, and the verification will not be successful in the final data verification. There are two solutions. Method 1: when entering the task of measuring temperature and humidity, the priority of this task is set to the highest; then, it will be lowered back to the original priority after measurement is completed. Method 2: when entering the task of temperature and humidity measurement, pause the task switching, then turn on the task switch after the measurement is completed. Of course, they will have a certain impact on the whole system. The most obvious consequence is that users will obviously feel that the response of the system is not timely. Finally, combined with their advantages and disadvantages, this system did not use these two methods, because the indoor temperature will not change very fast, and it only needs another 30 seconds to have a successful test to meet the actual needs [21]. The flow chart of temperature and humidity measurement and display is shown in Figure 3.

3.3.3. Realization of ZigBee Search Nearby Node Information Task. When users want to control a node, they first need to retrieve the existence of the node nearby. Before ZigBee search nodes, some parameters of ZigBee need to be set such as channel of the network, PANid of the network, and the source address of target address. The design sets 2.4 G of network channel, 11 network channels, 1010 network IDs, and 2020 source addresses, in which the target address can be dynamically switched according to the control target. After configuring the basic information, search constantly the number of nearby nodes and their address information in the task and then send the received number of nodes and node address information to the interface layer for processing. It is worth noting that ZigBee is also a peripheral resource. When a task is using ZigBee to send messages, the ZigBee module needs to be protected. At this time, another task can no longer search the information of nearby nodes through the ZigBee module. So it is generally to add a lock to lock to the ZigBee module and prevent other tasks from accessing it to prevent competition. The flow of search for nearby node information is shown in Figure 4.

3.3.4. Realization of Node Information Display Task. After searching the nearby nodes, it needs to be displayed on the control interface to control them. According to the row comparison between the node currently searched and the node last displayed, if there is a node in the last display list but not in the current search, it will be deleted from the display list. If there is no node in the last display list but there is a node in the current search, it will be added to the display list for display [22]. It is displayed in a list box control of emWin. The flow of showing current node information is shown in Figure 5.

3.3.5. Receive Interface Layer Sends Out the Control Status Task of Light Control Node. The purpose of this task is to receive the light control status information sent by the interface layer in real time. Concentrate the control status here and prepare to send it to the light control node, for example, switch, node selection, light brightness, light intensity, and so on. The priority is high, and the task level is 4. The receiving of light control status information is mainly in the form of a mailbox. First, create a mailbox data type that controls the status information type and set the number of mail. Finally, initialize and allocate memory space. It should be noted whether the sender has blocked sending, that is, whether he has been waiting for the successful sending and has been received by the receiver, or just send without waiting. The receiver is similar. This design uses sender blocking, receiver not blocking, and receiver real-time view receiving mailbox, if there is a mailbox to receive, not ready to receive the next time. The flow of receive control command and send to node is shown in Figure 6.

3.3.6. Modifying Node Name Task. Because the node is displayed in the control panel in the form of the network address of the node, and the network address is in the form
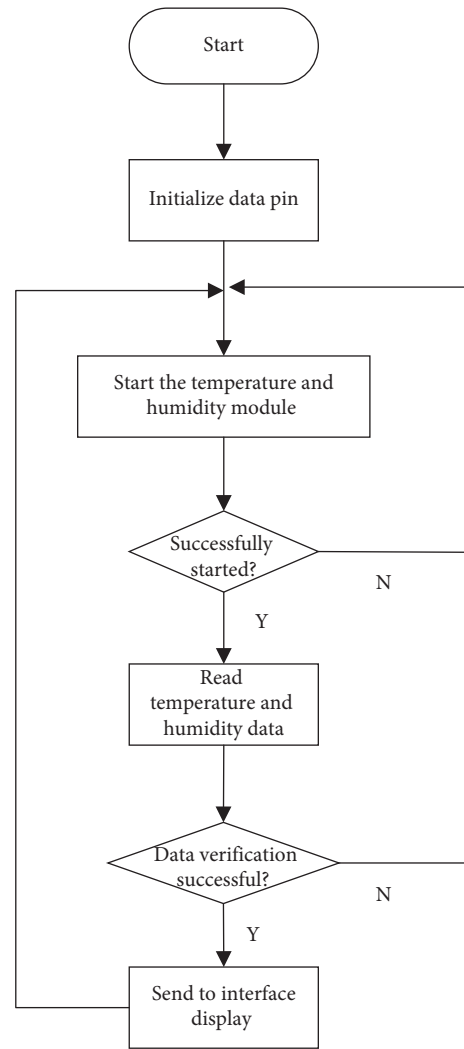
Figure 3: Flow chart of temperature and humidity measurement and display.
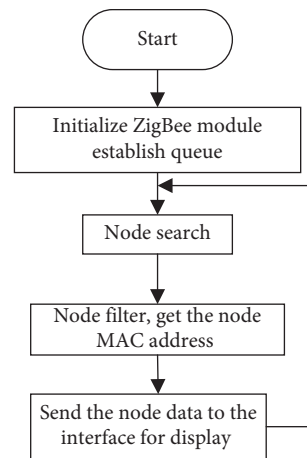
Figure 4: Flow of searching for nearby node information.

of a 16-byte number, the node that may be seen on the interface is displayed in the form of a hexadecimal byte similar to 02030507, which will be very inconvenient for
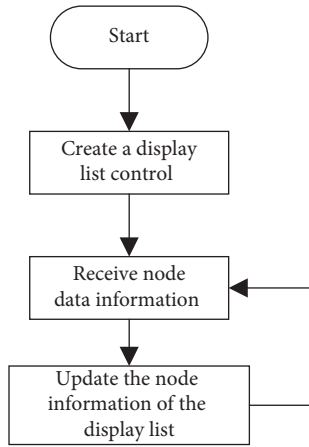
Figure 5: Flow of showing current node information.
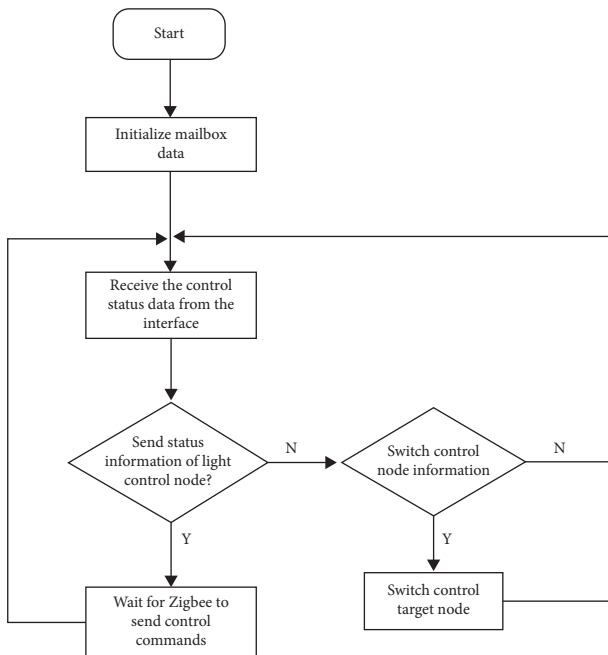
Figure 6: Flow of receive control command and send to node.

customers to use. Do not recognize where this node is or what special function it has from its name. In order to facilitate the use, set the name of the node that users can understand such as the living room and bedroom so that you easily identify and distinguish nodes. The flow of renaming node names is shown in Figure 7.

3.4. Data Layer Design of Light Control Node

3.4.1. Function Overview. The light control node is the controlled terminal equipment which is placed in any room or outdoor and controlled through the control panel to achieve the corresponding functions. The main functions are light on and off, light gradient, adjustable frequency flicker of light, timing switch, scene effect, brightness adjustment of light, and the color setting of light. The light control node structure is shown in Figure 8.
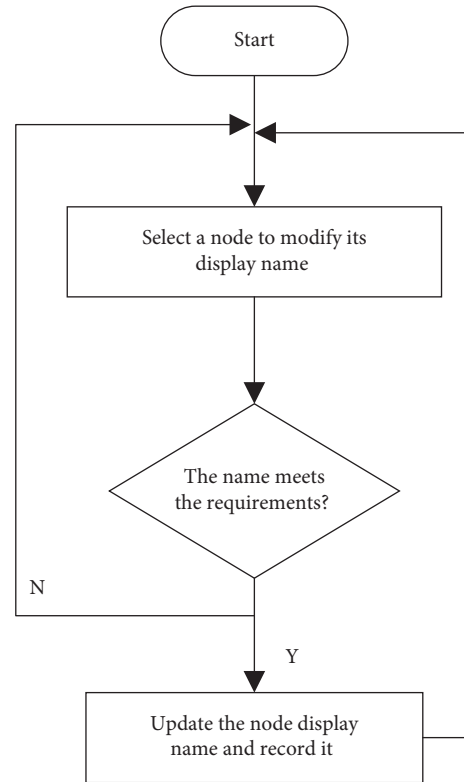
Figure 7: Flow of renaming node names.

3.4.2. Function Realization of Light Control Node. The software design of the light control node is like an infinite cycle state machine. Although it does not have the support of the operating system, it can continuously use the ZigBee module to receive the control signals from the control panel. Then, the control signal is parsed into the corresponding control command, entering the status query. Then execute the corresponding state function, that is, each function module that the light control node can achieve. These functions are to be debugged separately and sorted and packaged after debugging [23]. State machine implementation block diagram of light control node is shown in Figure 9.

4. System Debugging and Optimization

4.1. Communication Debugging between Multitasks. When programming in a system without an operating system, two different functions also need to be exchanged generally through the way of global variables to exchange data. Generally, data exchange is carried out through global variables. When ZigBee searches the current number of surrounding nodes and information, it needs to send the information to the interface task for display. Before sending, the data will be printed out by serial port printing. After receiving the data, the interface will print out the data which is always 0. The problem is that the memory here is cleared after sending which causes the received data to be always 0. Finally, the memory space will be cleared after receiving the data.
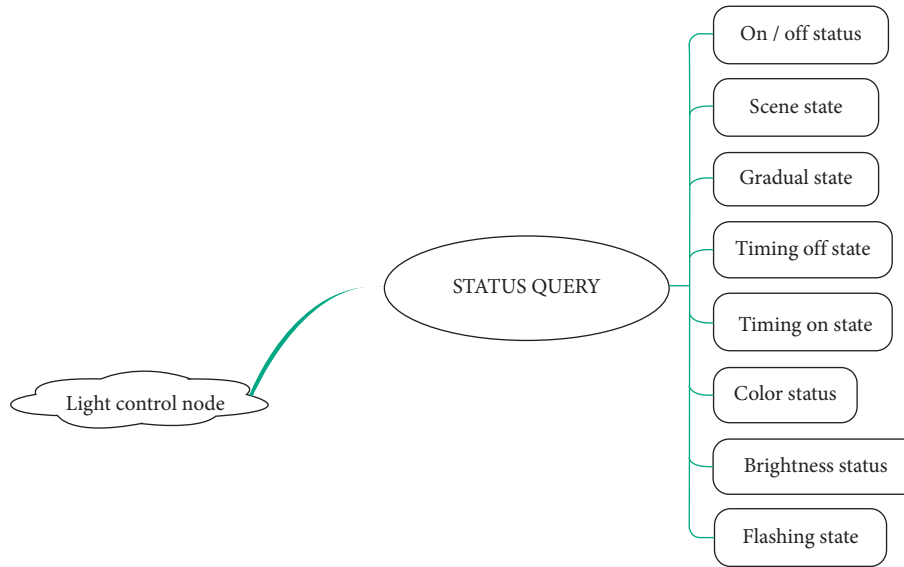
Figure 8: Light control node structure.
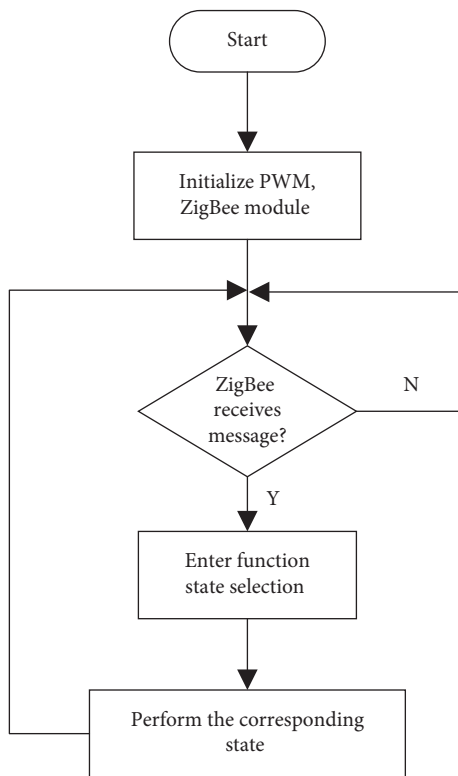
Figure 9: Light control node state machine.

In the process of multitasking, when different tasks use the same hardware device at the same time, it needs to be protected artificially. The most common method is to lock the module or memory and unlock it after accessing it. When the ZigBee module searches for nearby nodes, due to the scheduling of tasks, it is possible to send ZigBee information. At this time, there will be two tasks accessing the ZigBee module at the same time which will seriously lead to the death of the task and slightly lead to errors in the search

or the sending of information which can not achieve the expected effect of the program. This module needs to be protected. When task A accesses this module, it will lock this module. At this time, if task A is scheduled to execute another task B, it will wait until task A's access ends when it also needs to access this module. After task A is unlocked, Task B can have the opportunity to access this module which avoids the problem of multiple tasks accessing the same module at the same time. During debugging, it is found that ZigBee sending task is not executed.

In the process of node display and name change debugging, two identical nodes will be displayed. Sometimes, when the node name is changed, it still displays the previous node name. The solution is to create two linked lists. Linked list A records the node information currently displayed; linked list B records which node names have been changed. First, the node searched is compared with list B. The name of the node is changed. The node searched is compared with list B and updated displayed. When the node name is modified, the information and name of the node are added to list B, and then the node in list A is deleted.

Because ZigBee belongs to wireless communication, there is instability. When the number of surrounding nodes remains unchanged, the result of the last search may be different from that of the next time. The interface will flash especially unstable in this way. This requires filtering. The first filtering scheme: search five times each time, and it is determined to exist if two or more times are found; otherwise, it is determined to not exist so that the nodes displayed on the interface are stable. But there is a problem; that is, the task has been searching, occupying ZigBee module resources; other tasks have to wait until the end of the search task to access which is not real-time enough. The second filtering scheme is to establish a queue, put the searched node at the end of the queue each time, delete the node information at the head, and move the queue members next time, that is, move the first one to the position of the second

queue member and do it in turn. The number of member nodes of the queue is four and remains unchanged. As long as there are two or more members with this node, it will be judged that this node exists. In this way, each search can make it displayed on the interface, the number of searches is reduced, and the time of occupying the ZigBee module is four times less than the previous scheme. There are more opportunities for other tasks to not only occupy the ZigBee module but also improve the real-time performance of the task, but the disadvantage is that it takes about 3 seconds to remove a node. But this has little influence and the second design scheme is finally selected based on the practice.

In order to ensure the correctness of multitask execution, the function of real-time temperature and humidity detection is added. It only needs to ensure that the temperature and humidity verification is successful once in one minute. Otherwise, it takes 40 times to read the temperature and humidity data, and the reading of each bit of the data needs to delay to judge whether it is high level or low level. When this delay occurs, the CPU will switch to other tasks to improve its utilization, but when the delay time expires, due to the priority of the task, the task did not switch back on time which finally led to the unsuccessful verification of the measured data.

4.2. Debug of Light Control Node

4.2.1. Dimming with PWM Wave.
Pulse width modulation (PWM) is the time of high level output in a cycle. The adjustment of the duty cycle is the modulation of the pulse width. The color brightness of different RGB of duty cycles is different. Different colors can be modulated by outputting different PWM waves to the three channels of red green blue. The PWM module is like a timer built in. First, a count upper limit value (cycle) is set; then, a time t is set. When the counter is less than t, it controls the output low level of IO port. When the time is higher than t, the output of IO port is controlled as high level. When the counter reaches the upper limit, it will time again. So repeatedly, IO port outputs rectangular waveform [24].

4.2.2. Light State Adjustment.
The adjustment of color is mainly realized through the three primary colors of red, green, and blue. The RGB lamp can be configured with 256 different colors which depend on the adjustment of three PWM waves that control red, green, and blue. When the user sends the color switching command on the color selection panel of the control panel, the lamp control node will realize the color preparation by controlling three PWM waves of red, green, and blue.

Brightness adjustment: only change the brightness of the light, and keep the original color unchanged. The duty cycle of red, green, and blue PWM is equal to the increase or decrease of the column.

The flashing frequency of light can be controlled by a software program. The principle of flicker is to continuously control the light on and off showing certain regularity. The delay time between on and off is different with different flashing frequencies.

The gradient effect is realized by sending corresponding instructions from the controller. The lamp control node will make the lamp produce the effect of breathing lamp, slowly turn on, and then slowly turn off, and so on when the controller sends the gradient command.

Scene Effects: three different scene effect modes. Scene Effect 1: red, green, and blue lights turn on and off alternately, and with the passage of time, the speed of alternation is faster and faster. Scene Effect 2: use a random generating function to configure the color to be random and change once per second. Scene Effect 3: the breathing effect of red, green, and blue lights. The scene effect of each scene control signal is changed once.

Timing switch: the timing control lamp is on and off and is timed by the control end. When the preset time is over, the controller sends the control information of the timing switch.

4.3. Optimization of Light Control Node Real-Time State Switching.
When debugging the light control node, the control panel keeps sending control commands, but the light control node does not switch state in real time. Instead, when one state is executed, it will receive another information and then execute it. When a sent state is executed for a long time (scene effect), if the control interface sends another state (color switching) during execution, but the light control interface does not execute the latest sent state (color switching), it will continue to wait until the last state (scene effect) is executed. This will make people feel as if the command sent by the user has not been executed. Because the ZigBee module does not support interrupt, it can only use software to improve the real-time performance of the execution state. When executing a task (Scene Effect 1, red green blue cycle switch), the red light is on, keep the current execution state, exit this state, and go back to see if there is a control command received; if not, continue to execute the previous state and then switch the light to green. At this time, if the new control information is received, the control will be executed. This greatly increases the real-time performance of task switching and can interrupt the current task state to execute another state.

5. Summary

This design mainly involves ZigBee wireless communication technology, task allocation, and management of the operating system, resource management, and communication between multiple tasks, emWin interface design, PWM wave control of RGB lamp status, serial port debugging, and j-link debugging skills. The system supports ZigBee protocol communication, supports the human-computer interface, controls the light on and off, and controls the light brightness and color, gradient, flashing frequency; supports timing switch, and scene mode. It supports the automatic addition and deletion of lamp control nodes, can display indoor temperature data in real time, can stably display node

information, can freely modify the name of node display, and identify nodes with the same network address. The design has a wide range of applications. It provides technical support for the application of ZigBee technology in the field of light control.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] X. Huang, S. Garg, J. Nie et al., "Toward efficient data trading in AI enabled reconfigurable wireless sensor network using contract and game theories," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 98–108, 2022.

[2] X. Shen, W. Shao, Z. Zhang, and P. Xu, "Hotel intelligent guidance system based on ZigBee Technology," *Microprocessors and Microsystems*, vol. 9, no. 77, pp. 61–75, 2020.

[3] O. H. Seob, S. S. Beom, and L. G. Toe, "OCF-to-ZigBee (O2Z) bridging technique and IoTivity-based implementation," *Internet of Things Journal*, vol. 8, no. 22, pp. 16418–16426, 2021.

[4] A. E. Basabi, J. He, S. M. Hashemi, and X. Xuan, "A ZigBee software defined network security," *International Journal on Network Security*, vol. 24, no. 1, pp. 11–19, 2022.

[5] G. Meirong, "Smart campus teaching system based on ZigBee wireless sensor network," *Alexandria Engineering Journal*, vol. 61, no. 4, pp. 2625–2635, 2022.

[6] P. Hossein, K. Sangdeh, P. Zeng, and Z. Huacheng, "Securing ZigBee communications against constant jamming attack using neural network," *Internet of Things Journal*, vol. 8, no. 6, pp. 4967-4968, 2021.

[7] L. Wang, H. Zhu, L. Meng, and M. Ma, "Fungal industry supervision system based on zigbee technology," *2021 International Conference on Applications and Techniques in Cyber Intelligence*, vol. 81, pp. 788–792, 2021.

[8] P. S. Shiju Kumar and A. Ramesh Babu, "Enhanced STR based geographic location in ZigBee wireless networks: a hybrid PSACO method," *International Journal of Electronics*, vol. 108, no. 5, pp. 739–757, 2021.

[9] X. Li, T. Xu, and L. Q. Tien, "Exploration of underground mechanical environment monitoring system according to zigbee wireless communication," *International Journal of Mechatronics and Applied Mechanics*, vol. 2, no. 8, pp. 136–143, 2020.

[10] A. S. Deese and J. Daum, "ZigBee wireless smart plug network with RSSI multi-lateration-based proximity estimation and parallelised machine learning capabilities for demand response," *IET Wireless Sensor Systems*, vol. 10, no. 6, pp. 283–291, 2020.

[11] L. J. Kong, "Design and realization of distributed wireless temperature detection system based on Zigbee technology," *Australian Journal of Electrical and Electronics Engineering*, vol. 17, no. 1, pp. 47–55, 2020.

[12] L. Zhao, "Analysis of portable mechanical environmental monitoring system under zigbee wireless network," *International Journal of Mechatronics and Applied Mechanics*, vol. 1, no. 8, pp. 197–203, 2020.

[13] X. Chen, "The intelligent street light control system for preventing heavy fog of expressway based on ZigBee," *Wireless Personal Communications*, vol. 121, no. 1, pp. 353–359, 2021.

[14] P. Li, W. Wang, and B. Liu, "Effect of indoor environment on signal transmission strength of ZigBee wireless network," *Bing Gong Xue Bao*, vol. 41, pp. 275–279, 2020.

[15] V. T. Truong, A. Nayyar, and S. Ahmad Lone, "System performance of wireless sensor network using LoRa-Zigbee hybrid communication," *Computers, Materials & Continua*, vol. 68, no. 2, pp. 1615–1635, 2021.

[16] D. O. L. F. Pinto, M. L. Tiago, and L. P. D. G. Da, "Development of a smart traffic light control system with real-time monitoring," *Internet of Things Journal*, vol. 8, no. 5, pp. 3384–3393, 2021.

[17] X.-Z. Zhang and L.-S. Liu, "Design of the classroom intelligent light control system based on ARM9," *Advances in Smart Vehicular Technology, Transportation, Communication and Applications*, vol. 86, pp. 143–151, 2018.

[18] M. Michele, P. Tommaso, and B. Luca, "A smart LED light control system for environmentally friendly buildings," *Lecture Notes in Electrical Engineering*, vol. 409, pp. 181–189, 2017.

[19] N. Kumar, S. S. Rahman, and N. Dhakad, "Fuzzy inference enabled deep reinforcement learning-based traffic light control for intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 4919–4928, 2021.

[20] S. Michal, "ZigBee phase shift measurement approach to mobile inspection robot indoor positioning techniques," *Diagnostyka*, vol. 19, no. 3, pp. 101–107, 2018.

[21] Z. Qin, Y. Sun, J. Hu, W. Zhou, and J. Liu, "Enhancing efficient link performance in ZigBee under cross-technology interference," *Mobile Networks and Applications*, vol. 25, no. 1, pp. 68–81, 2020.

[22] L. Sheng, Y. Yang, and J. Zhang, "Wearable wireless calling system in hospitals based on ZigBee," *International Journal of Performability Engineering*, vol. 14, no. 8, pp. 1871–1876, 2018.

[23] Z. Jiaqi, B. Liu, and Y. Wang, "An improved method for indoor positioning based on ZigBee technique," *International Journal of Embedded Systems*, vol. 13, no. 3, pp. 292–299, 2020.

[24] F. Zhang, W. Jiang, Q. Lin, H. Wu, and S. Ning, "Remote and real-time monitoring system for indoor environment based on ZigBee and LabVIEW," *Sensors and Materials*, vol. 32, no. 6, pp. 2139–2158, 2020.

Research Article

A Small Network MicronNet-BF of Traffic Sign Classification

Hai-Feng Fang [ID],[1] **Jin Cao** [ID],[1] **and Zhi-Yuan Li** [ID][2]

[1]*School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang 212100, China*
[2]*College of Engineering and Computer Science, Australia National University, Canberra,*
 Australian Capital Territory 2601, Australia

Correspondence should be addressed to Zhi-Yuan Li; zhiyuan.li1@anu.edu.au

One of a very significant computer vision task in many real-world applications is traffic sign recognition. With the development of deep neural networks, state-of-art performance traffic sign recognition has been provided in recent five years. Getting very high accuracy in object classification is not a dream any more. However, one of the key challenges is becoming making the deep neural network suitable for an embedded system. As a result, a small neural network with as less parameters as possible and high accuracy needs to be explored. In this paper, the MicronNet which is a small but powerful convolutional neural network is improved by batch normalization and factorization, and the proposed MicronNet-BN-Factorization (MicronNet-BF) takes advantages about reducing parameters and improving accuracy. The effect of image brightness is reduced for feature recognition by the elimination of mean and variance of each input layer in MicronNet via BN. A lower number of parameters are realized with the replacement of convolutional layers in MicronNet, which is the inspiration of factorization. In addition, data augmentation is also been changed to get higher accuracy. Most important, the experiment shows that the accuracy of MicronNet-BF is 99.383% on German traffic sign recognition benchmark (GTSRB) which is much higher than the original MicronNet (98.9%), and the most influence factor is batch normalization after the confirmation of orthogonal experimental. Furthermore, the handsome training efficiency and generality of MicronNet-BF indicate the wide application in embedded scenarios.

1. Introduction

Traffic signs, usually erected at the side of roads, use texts or symbols to provide road information for vehicle divers and pedestrians (see Figure 1). Traffic sign recognition is essential in advanced driver assistance systems (ADASs) and autonomous vehicles [1]. In the real cases, camera installed on the vehicle takes photos of roads. The information processing system processes the image and detect and classify the traffic sign according to its characters. The classification result provides road information for drivers or adjusts the motion state of an autonomous vehicle. Because the captured images are affected by brightness and weather conditions, traffic sign classification has high requirement in accuracy and robustness.

For the sufficient research of traffic sign recognition, researchers established a multitraffic sign recognition dataset such as German Traffic Sign Recognition Benchmark

(GTSRB) [2], Belgium Traffic Sign Classification Benchmark (BelgiumTSC) [3], and Tsinghua-Tencent 100K dataset [4]. GTSRB dataset provides 51,840 colorful images of German road signs in 43 classes. This dataset also provides cropped images for accurate classification. Most of images are clear, but part of them is blurred and darkness used to test the algorithm's robustness. It not only allows researchers to test the accuracy of their algorithm and to compare it with human performance but also to be transformed by the histogram of the oriented gradient algorithm to prevent projection distortion [5] or denoised to promise the quality of dataset [6].

In recent years, convolution neural networks (CNNs) show high performance in the GTSRB dataset [7–9]. CNNs, inspired from human's visual perception mechanism, are applied broadly in computer vision [10]. As a deep learning network, it has many layers to simulate neurons to learn the characters of images. It has showed high performance in

FIGURE 1: Various categories of traffic signs under different weather and brightness in GTSRB.

many datasets such as CIFAR [11] and ImageNet [12], so people consider applying the enhanced CNN (e.g., LeNet-5 [13], Caps Net [5], PFANet [14], differential evolution evolved RBFNN [15], etc.) in traffic classification. However, application in vehicles has its restriction. The network requires high response speed under the limited storage space. The hardware installed on the vehicle does not have enough computation ability, which causes the scale of the network limited [16]. As such, some famous mature deep networks such as GoogLeNet [17] and VGG [18] are too deep or huge to be applied in vehicles directly. However, small networks are feasible. Zhang et al. [1] proposed two light weight CNN simple student network and deep teacher network and assisted the training of the student model to achieve high accuracy in traffic sign classification. Arman et al. [19] proposed a novel thin yet deep convolutional neural network for a light weight architecture. Cao et al. [20] used HSV color space preprocess images and applied improved LeNet-5 CNN model with a small number of parameters in traffic sign classification. Although the research studies above have slimmed the networks to adapt the embedded system, the recognition of brightness and blur pictures of traffic sign is still an arduous challenge. Wong et al. [21] proposed MicronNet and trained it with the augmented (e.g., HSV augmentation, Gaussian blur, motion blur, etc.) traffic sign photos. However, the augmented dataset in [21] have no emphasis, causing information redundancy. There are still a number of optimizations in the structure of MicronNet and the augmentation of traffic sign dataset, and it is essential to find a proper balance between the processing of a traffic sign

dataset and the light weight structure of the convolutional neural network.

Inspired from the above network, we proposed a CNN based on MicronNet, a small network and overcomes drawbacks of the original network. In this study, we mainly focus on MicronNet-BN-Factorization (MicronNet-BF) which fused the superiority of MicronNet, batch normalization, and factorization. In addition, the appropriate augmentation methods of insufficient illumination traffic signs are selected for a better training performance.

The main contributions of this paper can be summarized as follows:

> The complicated data augmentation methods (including shift, flip, mirror, HSV, blur, and rotation) are simplified into shift, scale, and the V channel of HSV, avoiding that too much data augmentation may introduce some useless or even false characters of traffic sign and reduce the accuracy of the neural network.

> Two channels are additionally supplemented to the first layer, 1-by-1 convolution, to enhance the learning of image features from the dataset. 5-by-5 convolutional layer is replaced by two sequential 3-by-3 convolutional layers, reducing parameters and extracting more meticulous characteristics to increase the accuracy.

> Batch normalization can learn and fix the input means and variances of each layer. For the traffic sign recognition, the adverse effect of brightness is effectively reduced, improving the classification accuracy of insufficient lighting images.

2. Related Work

MicronNet [21], a small deep convolutional neural network, is proposed to achieve real-time embedded traffic sign classification. The network structure is optimized from a large network by repeating omitting parameters and testing network to maintain high accuracy with the least number of useful parameters. The final optimized network reaches 98.9% accuracy only containing 0.51 M parameters and which is competitively with the deep inception-based CNN [22] with 10.5 M or single CNN with 3 STNs [23] with 14 M, etc. Furthermore, a few logical operations are required for MicronNet to perform inferences and short computation time meanwhile. However, the network cannot deal with dark and blurred images well (see Figure 2). Based on the MicronNet, we adjust data augmentation and modify parts of the network to make it suitable to both common images and dark images.

Ioffe and Szegedy [24] proposed batch normalization (BN) to improve classification accuracy and training rate. Because of internal covariate shift, the changed parameters of the previous layers causing each layer inputs changed after every training epoque, and traditional network training chooses a low learning rate. Batch normalization normalize every layer's input for each training minibatch. We introduce batch normalization to MicronNet to improve its learning rate and accuracy.

Szegedy et al. [25] based on GoogLeNet [17] presented the inception V2 network. In this paper, Szegedy presented a theory that two sequential small convolutional filters can replace a large convolutional filter to improve the learning rate and reduce parameter number and achieve similar accuracy because the receptive field of two methods are the same. The factorizing mentality is fused in MicronNet either.

3. Proposed Methods

3.1. Data Augmentation. The uneven distribution of data will be decreasing the accuracy of classification. Researchers use a various of data augmentation techniques to balance the number of samples [21, 26]. However, on the one hand, the data that can be augmented based on one sample is limited and cannot be increased indefinitely due to the distortion of the characteristics of sample in the process of augmentation. On the other hand, the ministructure of the neural network cannot effectively learn too much characteristics. As a result, the proposed data augmentation is simplified to three ways: (i) shifting, (ii) brightness, and (iii) scale. The properties of choosing these three ways can be described as follows: (1) Shifting can help to deal with the partially covered traffic signs. (2) Brightness can help to learn the traffic signs under different light conditions. (3) Scale can help to handle various sizes of traffic signs. The examples can be seen in Figure 3.

3.2. MicronNet-BF. MicronNet is a compact deep neural network proposed for traffic sign classification on embedded devices [21]. It has struck a relative balance between the

augmentation of a traffic sign dataset and the simplifying of the network architecture, but the main problem in the example of misclassified traffic images is either heavily motion blurred (left), partially occluded (middle), or exhibit poor illumination (right). Based on the MicronNet and inspired from the network architecture of inception V2 [21, 25], an improved network architecture MicronNet-BN-factorization (MicronNet-BF) is proposed in this paper. MicronNet-BF is taken to (1) improve the total accuracy on traffic sign recognition problems, (2) keep the same model size or achieve a smaller model size for embedded devices, (3) achieve better performance on classification accuracy of a special class (low brightness images).

Figure 4 shows the overall network architecture of MicronNet-BF, and Table 1 prints the details of parameters. In this architecture, it mainly has 5 convolutional layers, 2 fully-connected layers, and a SoftMax layer. All the activation functions in this network are chosen to be rectified linear unit (ReLU) for the reducing of computational complexity. In this network, the 1-by-1 convolutional layer in the original MicronNet is extended to have 3 output channels, and the 5-by-5 convolutional is replaced by two of the 3-by-3 layers. Furthermore, batch normalization layers are added into the proposed network to deal with the brightness difference in the input images and improve training speed.

The batch normalization layer in a network learns the mean and variance of dataset, and fixes the input means and variances of each layer. In the application of traffic sign classification, the brightness of each input image is closely related to the mean and variance value of the image. By normalization of the mean and variance of each image, the batch normalization layer turns all the images in the dataset to have a similar brightness, which improves the classification accuracy of the low brightness traffic sign images.

On the one hand, inspired by the idea of "factorization" into smaller convolutions in inception V2 [25], the 5-by-5 convolutional layer is replaced by two of the 3-by-3 convolutional layers, as shown in Table 1. The 3-by-3 convolutional layer used in this replacement enables the network to learn some smaller scaled feature from the input images and share the features among the following up 3-by-3 convolutional layer. Furthermore, the spatial coverage of the original 5-by-5 layer is maintained by the overlap of two 3-by-3 convolutional layers. In this way, this improvement results in a slight deeper network with the ability of learning smaller scaled details from the traffic signs, which significantly improved the overall classification accuracy.

On the other hand, traffic signs are normally designed with colors of high contrast, including black, white, red, yellow, and blue. In order to use the color information in the traffic sign classification, the 1-by-1 convolutional layer is extended to have 3 output channels. In the traditional network, the 1-by-1 convolutional layer combines the RGB color of the input image to 1 value on each pixel location, which can be considered as a RGB to gray conversion. After extending the output channel to 3, the 1-by-1 convolutional layer becomes a color extraction layer, which provides 3 different color combinations for the following up layers.

FIGURE 2: Examples of traffic images in GTSRB unrecognized by original MicronNet since blur or darkness.

FIGURE 3: Data augmentation of traffic signs in GTSRB. The data augmentation ways about shifting, brightness and scale can be seen from images.

4. Experimental Evaluation

4.1. Implementation Details. For a fair comparison, both methods and algorithm architectures take similar learning hyperparameters. A learning rate of 0.007 is used for MicronNet and MicronNet-BF in most scenes. The hyperparameters of comparison methods are taken as the default of the corresponding cite paper. Meanwhile, the same experiment platform is used in all experiment and is a Linux Ubuntu20.4 operating system with PyTorch. The GPU is GTX1080ti of NVIDIA. As an evaluation index, the accuracy rate refers to the percentage of the number of correctly recognized in the test dataset to the whole number of test dataset, and time refers to the sum of training time of 100
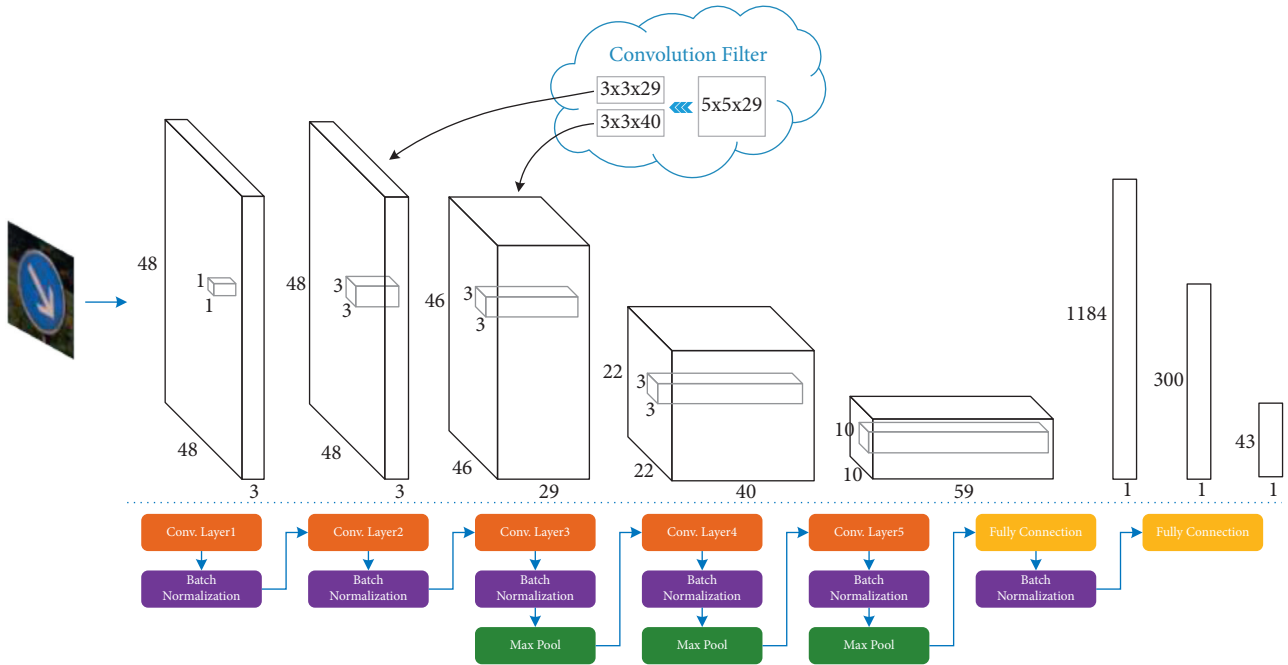
FIGURE 4: Brief structure diagram of MicronNet-BF. The sizes of convolution filters and convolution layer's inputs are expressed on the upper diagram, and the brief description of MicronNet-BF's complete layer structure on the below.

TABLE 1: The structure of proposed MicronNet-BF.

Type/stride/pad	Filter shape	Input size
Conv/s1/p0	$1 \times 1 \times 3$	48×48
BN	Batch normalization	48×48
Conv/s1/p0	$\mathbf{3 \times 3 \times 29}$	48×48
BN	Batch normalization	46×46
Conv/s1/p0	$\mathbf{3 \times 3 \times 40}$	46×46
BN	Batch normalization	44×44
Pool/s2/p0	3×3	Maxpool
Conv/s1/p0	$3 \times 3 \times 59$	22×22
BN	Batch normalization	22×22
Pool/s2/p0	3×3	Maxpool
Conv/s1/p0	$3 \times 3 \times 74$	10×10
BN	Batch normalization	10×10
Pool/s2/p0	3×3	Maxpool
FC/s1	1×300	1×1184
BN	Batch normalization	1×300
FC/s1	1×300	1×300
Softmax/s1	Classifier	1×43

The bold numbers mean one of the innovations in applying inception V2 to MicronNet.

TABLE 2: Abbreviations of networks.

Abbreviation	Network
MicN	MicronNet (self-trained)
MicN-B	MicronNet + BN
MicN-O	MicronNet + 1-by-1 × 3 outputs
MicN-3	MicronNet + 2 × 3-by-3
MicN-L	MicronNet + insufficient lighting
MicN-BF	MicronNet-BN-factorization (the fusion of all above)
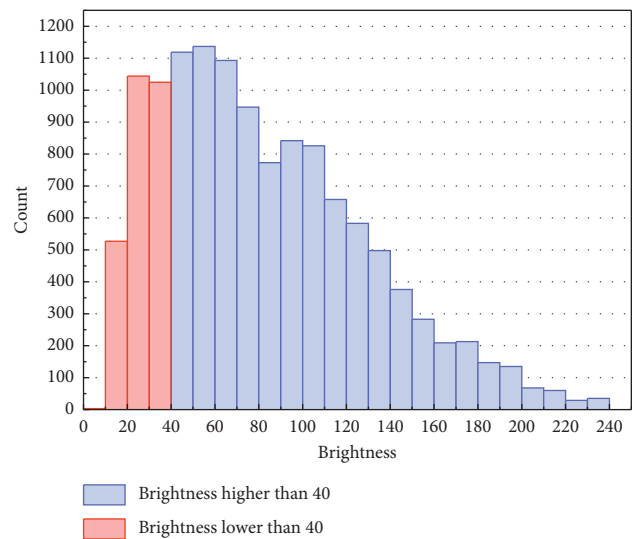
FIGURE 5: Testing set brightness distribution. The samples with brightness lower than 40 are constituted to the brightness insufficient dataset.

samples in each epoch. Additionally, some abbreviations for networks have been adopted for briefly expression, as shown in Table 2.

For the comparing of the recognition of networks on dark images, the insufficient brightness images of traffic sign are extracted from the testing dataset to combine a new challenging dataset. After ordering the brightness of the whole testing dataset, the first 20.57% samples (the number of 2599) were used as the new insufficient illuminated traffic sign dataset; that is, the average brightness of each sample in the new dataset was lower than 40. The quantity distribution of the testing dataset is shown in Figure 5. The samples with

red are constructed to a harder dataset, and the rest samples are used for testing.

4.2. MicronNet-BF Evaluation with GTSRB.

For the comparing with MicronNet [21], the proposed MicronNet-BF is evaluated on German traffic sign recognition benchmark (GTSRB) [27] firstly. The GTSRB dataset contains color traffic sign images from 43 classes and intends for recognition. On the one hand, the evaluation with the overall accuracy on GTSRB is taken normally. For further challenges, the recognition of lower brightness images from the GTSRB testing set is processed meanwhile. On the other hand, during the training of the network, rotation, shifting, and scaling are used as data argumentation strategies to improve the generality of the resulting network, especially for the testing images with partly visible sign.

Figure 6 shows the testing accuracy and training time of the proposed network MicronNet and the comparison networks based on GTSRB dataset. Comparing with the MicronNet, the batch normalization layer added into it improves the classification accuracy from 97.686% to 98.74%. Furthermore, the extending of output channels on 1-by-1 layer improves the accuracy to 97.561%, and the replacement of two 3-by-3 layers improves the overall accuracy to 98.777%. Thus, the overall accuracy with 99.383% of the MicronNet-BF is improved by the three strategies proposed in this project, respectively. What is more, the comparison of MicN-BF × L with 99.448% and MicN-L with 98.147% indicates the great recognition performance of MicronNet-BF under insufficient brightness.

4.3. Validation of MicronNet-BF Influence Factors.

In the front section experiment, it was proved that batch normalization, extending of output channels on 1-by-1 layer, and factorization were successfully integrated into MicronNet, but the effect and influence processes of each factor need further experimental verification. The Taguchi orthogonal array experimental method can greatly reduce the number of experiments than grid searching experiment and inference of the optimal parameter combination by the orthogonal method [28, 29]. The Taguchi orthogonal array experimental method is used to obtain the optimal values and evaluate the influence of factors [30, 31].

There are three factors in MicronNet-BF that need to be focused on. In addition, the interaction between factors should also be considered, including dataset with insufficient illumination. For a summary, there are seven factors (including MicN-B, MicN-O, MicN-3, MicN-L, MicN-B × O, MicN-B × 3, and MicN-O × 3) and two levels (including N (NULL) and A (APPLY)) in the experiment. With respect to selected factors, $L_8(2^7)$ orthogonal array was designed for the validation experiments, and the experimental design and the results are shown in Table 3.

For an immediate point of view, the best accuracy with 99.448% is taken with the network of MicronNet-BF under the insufficient brightness testing dataset; it is consistent with the conclusion of the previous subsection. In Table 3, I_{A1} denotes the summary of accuracy under the first level of the factors, and I_{A2} for the second level, R_A represents the absolute value of the difference of I_{A1} and I_{A2}. The meaning of I_{T1}, I_{T2}, and R_T is similar with the third before but for time. From the row of R_A, the biggest difference is the factor of MicN-B and the smallest is MicN-O, and it indicates that the factor MicN-B has the most influence of the recognition accuracy of traffic signs, and the factor MicN-O has the lowest influence. The result in the row of RT shows that the factor MicN-B has the most influence of the training time too, but the factor MicN-B X O has the lowest. According to the difference value of interaction factor, there is only a tiny effect about accuracy and time. Therefore, the ranking of effects can be sorted as MicN-B > MicN-3 > MicN-L > MicN-B × 3 > MicN-O × 3 > MicN-B X O > MicN-O.

For the insufficient brightness traffic sign dataset, the recognition accuracy is shown in Figure 7. The best accuracy with 99.448% is taken by MicronNet-BF, and the accuracies of MicN-B × L with 98.936% and MicN-3 × L with 99.079% are better than MicronNet significantly. The recognition ability of batch normalization and factorization for traffic signs with insufficient brightness is proved. Although the accuracy of MicN-O × L with 97.96% is no better than others, this tendency can be seen in Table 4 with the R_A of MicN-O. On the one hand, it is indicated that extending the output of the 1 by 1 layer to three channels cannot enhance the recognition performance of traffic signs with insufficient brightness, but MicN-O can improve the classification performance of traffic signs with normal illumination and rich colors to a certain extent, and hardly increase the extra training time meanwhile. Therefore, the MicN-O is also effective.

On the other hand, the fluctuation of loss value and accuracy rate in the process of iteration can also reflect the role of various factors. As Figure 8 shows, networks present various fluctuation trends in the training process. MicN fluctuated widely in the first 10 iterations and remained fairly flat thereafter. The loss value of MicN-B dropped quickly, but the subsequent fluctuations lasted for a long time. The loss value of MicN-O dropped faster than MicN and have a bit fluctuation later. MicN-3 get the best performance in the process of iteration, dropped fastest, and more flatted. Finally, under the balance of various factors, MicN-BF loss value decreases rapidly with few fluctuations so as to achieve the best classification performance quickly and maintain stability.

4.4. Comparison Evaluation.

With the discussion in the previous subsections, the test of MicronNet-BF on GTSRB dataset is quite complete. In order to further verify the recognition performance of MicronNet-BF on addition traffic sign dataset and different types of datasets, some representative datasets were selected. The properties of several dataset and the evaluation performance are listed in Table 4.

The Belgium Traffic Signs Classification dataset has 62 categories, 4,591 samples for training and 2498 for testing. The results show that the recognition performance of MicronNet-BF with 82.122% is better than MicronNet with
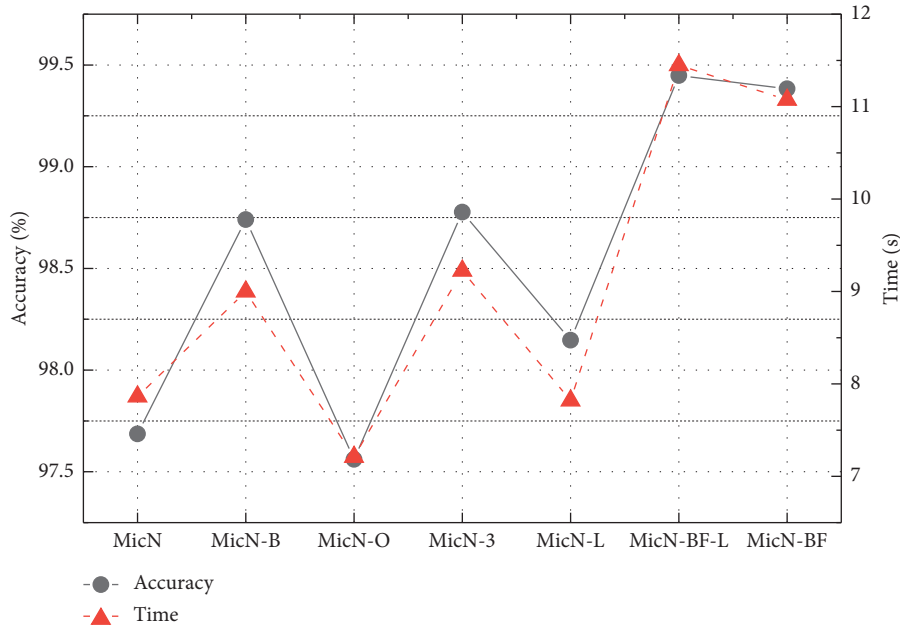
FIGURE 6: Accuracies and time of networks based on GTSRB. The best accuracy with 99.448% is taken with MicronNet-BF under the insufficient brightness dataset.

TABLE 3: Orthogonal experimental design and the results.

Ex. No.	MicN-B	MicN-O	MicN-B × O	MicN-3	MicN-B × 3	MicN-O × 3	MicN-L	Acc (%)	Time (s)
1	N	N	N	N	N	N	N	97.686	**7.863**
2	N	N	N	A	A	A	A	99.079	9.564
3	N	A	A	N	N	A	A	97.96	8.371
4	N	A	A	A	A	N	N	98.534	8.858
5	A	N	A	N	A	N	A	98.936	9.708
6	A	N	A	A	N	A	N	99.28	11.707
7	A	A	N	N	A	A	N	98.824	9.509
8	A	A	N	A	N	N	A	**99.448**	11.448
I_{A1}	393.259	394.981	395.037	393.406	394.374	394.604	394.324		
I_{A2}	396.488	394.766	394.71	396.341	395.373	395.143	395.423		
R_A	**3.229**	**0.215**	0.327	2.935	0.999	0.539	1.099		
I_{T1}	34.656	38.842	38.384	35.451	39.389	37.877	37.937		
I_{T2}	42.372	38.186	38.644	41.577	37.639	39.151	39.091		
R_T	**7.716**	0.656	**0.26**	6.126	1.75	1.274	1.154		

The bold numbers represent the best performance of the corresponding item.
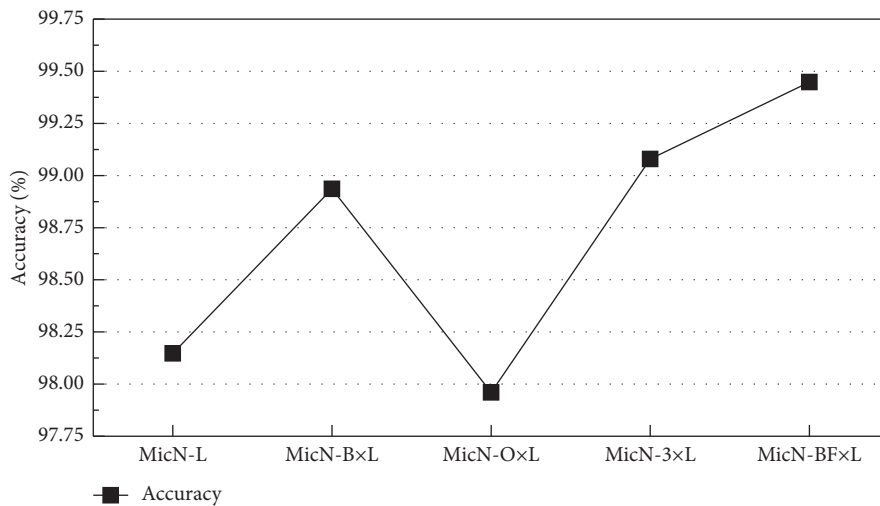
FIGURE 7: Accuracies of networks based on insufficient brightness data. The best accuracy with 99.448% is taken by MicronNet-BF.

TABLE 4: The testing accuracy and training time.

	Dataset	GTSRB	BTSC [3]	MNIST	SVHN [32]	Cifar10	Cifar100
Property	Categories	43	62	10	10	10	100
	Training samples	39209	4591	60000	73257	50000	50000
	Testing samples	12630	2498	10000	26032	10000	10000
MicronNet	Testing accuracy	97.686%	80.388%	99.49%	92.16%	34.829%	10.333%
	Training time	1 s	1.01 s	0.55 s	0.38 s	0.85 s	0.85 s
MicronNet-BF	Testing accuracy	**99.383%**	**82.122%**	**99.58%**	**93.35%**	**78.67%**	**49.93%**
	Training time	1.41 s	1.48 s	0.60 s	0.65 s	1.29 s	1.27 s

The bold numbers represent the one with the better test accuracy in MicronNet-BF and MicronNet.
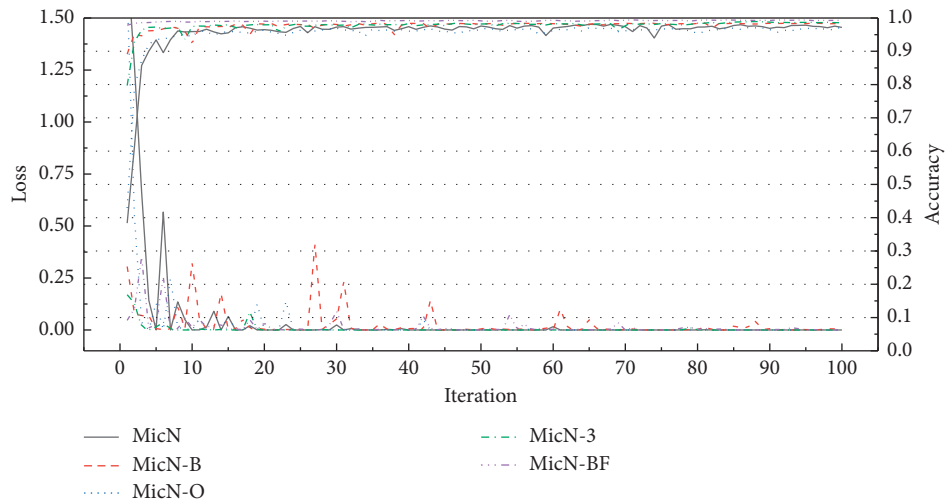
FIGURE 8: The accuracies and losses of networks. Obviously, the loss value of MicN-B is more volatile.

TABLE 5: Classification accuracies of GTSRB and parameter number of algorithms.

Network	Parameter (M)	Accuracy (%)
Deep inception-based CNN [24]	10.5	99.81
Single CNN with 3 STNs [30]	14	99.71
HLSGD [33]	23.2	99.65
Student network [1]	0.73	99.61
MCDNN [16]	38.5	99.46
MicronNet-BF	**0.44**	**99.38**
MicronNet [21]	0.55	98.9
DNN [16]	1.54	98.52

The bold characters and numbers represent the proposed algorithm and its parameters, respectively.

80.388%. It indicated that with the training of a small number of traffic signs, the classification performance of MicronNet-BF decreased, but it was still higher than MicronNet. On the one hand, the generalization of MicronNet-BF and MicronNet has been verified with the accuracies of 99.58% and 99.49% on the dataset of MNIST. In the more challenging number classification dataset of SVHN, MicronNet-BF maintained a slight advantage, indicating that the structural superiority is not limited to the recognition of traffic signs. In the case of more complex dataset Cifar10 and Cifar100, MicronNet-BF is unable to learn deeper features as to its lightweight structure, and the

recognition accuracy is only 78.67% and 49.93%, respectively, but it still far exceeds MicronNet with 34.83% and 10.33%.

On the other hand, the MicronNet-BF is mainly used in embedded devices. By comparing the difference of structure between MicronNet-BF and MicronNet, replacing a 5-by-5 convolution filter with two 3-by-3 convolution filters has the greatest impact on the number of variables, and the number of variables is reduced to 0.44 M. As listed in Table 5, compared with the state-of-art networks, with the minimum number of variables of 0.44 M, MicronNet-BF has achieved excellent results with a difference of no more than 0.4% compared with the larger networks.

5. Conclusions

In order to improve the recognition performance of traffic signs further, the MicronNet-BF which is fused by MicronNet, batch normalization, and factorization is proposed. The addition of batch normalization enhances the recognition performance to 98.74%, which is 1.05% higher than the performance of MicronNet. The application of factorization improves the accuracy to 98.77%. The MicronNet-BF which is combined by multifactors listed above has a recognition performance with 99.383% which has a great improvement than MicronNet. On the one hand, the batch

normalization and factorization do enhance the ability of recognizing the traffic signs with insufficient brightness after the experiment evaluation. On the other hand, the most influence factor is batch normalization after the confirmation of orthogonal experimental. In the end, the performance of MicronNet-BF used in BTSC, Cifar10, and Cifar100 is better than MicronNet.

Although the algorithm is applied in the embedded system, the less parameters are not better, and striking a balance between the number of parameters and the size of storage space needs further study.

Data Availability

All the dataset could be found in https://paperswithcode.com/datasets, including GTSRB, BTSC, Mnist, SVHN, Cifar10, and Cifar100.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] J. Zhang, W. Wang, C. Lu, J. Wang, and A. K. Sangaiah, "Lightweight deep network for traffic sign classification," *Annals of Telecommunications*, vol. 75, no. 7, pp. 369–379, 2020.

[2] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural Networks*, vol. 32, pp. 323–332, 2012.

[3] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3d localization," *Machine Vision and Applications*, vol. 25, no. 3, pp. 633–647, 2014.

[4] Z. Zhe, D. Liang, S. Zhang, X. Huang, and S. Hu, "Traffic sign detection and classification in the wild," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2110–2118, Las Vegas, NV, USA, June 2016.

[5] S. He, L. Chen, S. Zhang et al., "Automatic recognition of traffic signs based on visual inspection," *IEEE ACCESS*, vol. 9, pp. 43253–43261, 2021.

[6] S. Liu, T. Cai, X. Tang, Y. Zhang, and C. Wang, "Visual recognition of traffic signs in natural scenes based on improved RetinaNet," *Entropy*, vol. 24, no. 1, p. 112, 2022.

[7] F. M. Johner and J. Wassner, "Efficient evolutionary architecture search for cnn optimization on gtsrb," in *Proceedings of the 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), December*, vol. 16-19, pp. 56–61, Boca Raton, FL, USA, December 2019.

[8] R. Qian, Y. Yue, F. Coenen, and B. Zhang, "Traffic sign recognition with convolutional neural network based on max pooling positions," in *Proceedings of the 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 578–582, Changsha, China, Aug 13-15.

[9] Z. Qiumei, T. Dan, and W. Fenghua, "Improved convolutional neural network based on fast exponentially linear unit activation function," *IEEE Access*, vol. 7, pp. 151359–151367, 2019.

[10] J. Gu, Z. Wang, J. Kuen et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.

[11] A. Krizhevsky and G. Hinton, *Learning Multiple Layers of Features From Tiny Images*, p. 280, University of Toronto, Toronto, Ontario, 2009.

[12] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "Imagenet: a large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Miami, FL, USA, June 2009.

[13] Z. Ameur, L. Anis, and S. Anis, "A lightweight model for traffic sign classification based on enhanced LeNet-5 network," *Journal of Sensors*, vol. 2021, Article ID 8870529, 13 pages, 2021.

[14] Z. Ke, Z. Yufei, and F. Dongmei, "Learning region-based Attention network for traffic sign recognition," *Sensors*, vol. 21, no. 3, pp. 1–19, 2021.

[15] R. Manasa, K. Karibasappa, and S. M. Kumar, "Differential evolution evolved RBFNN based automated recognition of traffic sign images," *Results in Control and Optimization*, vol. 5, pp. 1–13, 2021.

[16] D. Cireş̧An, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Networks*, vol. 32, pp. 333–338, 2012.

[17] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, Boston, MA,USA, June 2015.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," p. 1556, 2014, https://arxiv.org/abs/1409.1556.

[19] H. W. A. Arman, S. Arefin, A. S. M. Shihavuddin, and M. A. Hasan, "DeepThin: a novel lightweight CNN architecture for traffic sign recognition without GPU requirements," *Expert Systems with Applications*, vol. 168, Article ID 114481, 2021.

[20] J. Cao, C. Song, S. Peng, F. Xiao, and S. Song, "Improved traffic sign detection and recognition algorithm for intelligent vehicles," *Sensors*, vol. 19, no. 18, Article ID 4021, 2019.

[21] A. Wong, M. J. Shafiee, and M. St Jules, "Micronnet: a highly compact deep convolutional neural network architecture for real-time embedded traffic sign classification," *IEEE Access*, vol. 6, pp. 59803–59810, 2018.

[22] M. Haloi, "Traffic sign classification using deep inception based convolutional networks," 2015, http://arxiv.org/abs/1511.02992.

[23] Á. Arcos-García, J. A. Álvarez-García, and L. M. Soria-Morillo, "Deep neural network for traffic sign recognition systems: an analysis of spatial transformers and stochastic optimisation methods," *Neural Networks*, vol. 99, pp. 158–165, 2018.

[24] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," pp. 448–456, 2015, http://arXiv.org/abs/1502.03167.

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and*

pattern recognition, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[26] D. Tabernik and D. Skoaj, "Deep learning for large-scale traffic-sign detection and recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 99, pp. 1–14, 2019.

[27] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The german traffic sign recognition bench mark: a multi-class classification competition," in *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 1453–1460, San Jose, CA, USA, Aug 2011.

[28] B. Cem and K. Tahsin, "Proper estimation of surface roughness using hybrid intelligence based on artificial neural network and genetic algorithm," *Journal of Manufacturing Processes*, vol. 70, pp. 560–569, 2021.

[29] R. R. Nasibeh, A-T. Mohammad-R, and A. Alireza, "Experiment-based affect heuristic using fuzzy rules and Taguchi statistical method for tuning complex systems," *Expert Systems with Applications*, vol. 172, Article ID 114638, 2021.

[30] K. Balaji, M. Siva Kumar, and N. Yuvaraj, "Multi objective taguchi–grey relational analysis and krill herd algorithm approaches to investigate the parametric optimization in abrasive water jet drilling of stainless steel," *Applied Soft Computing*, vol. 102, Article ID 107075, 2021.

[31] L. Zhang and M. Zhang, "Image reconstruction of electrical capacitance tomography based on optimal simulated annealing algorithm using orthogonal test method," *Flow Measurement and Instrumentation*, vol. 80, Article ID 101996, 2021.

[32] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and Y. Andrew, "Reading digits in natural images with unsupervised feature learning," in *Proceedings of the NIPS Workshop on Deep Learning and Unsupervised Feature Learning*, December 2011.

[33] J. Jin, K. Fu, and C. Zhang, "Traffic sign recognition with hinge loss trained convolutional neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 1991–2000, 2014.

Research Article

IoT Architecture-Based Mechanism for Digital Transmission of Key Aspects of the Enterprise

Cheng Chen (ID)

School of Economics and Management, Tongji University, Shanghai 200092, China

Correspondence should be addressed to Cheng Chen; ccacadia@tongji.edu.cn

This paper provides an in-depth study and analysis of data transmission in key aspects of the enterprise using internet of things (IoT) technology. IoT technology can provide an effective solution to the integration of enterprise resources and efficiency improvement. If it can be properly introduced into the enterprise, it not only can effectively integrate the enterprise resources in the management but also can significantly improve the efficiency and thus further reduce its operating cost. This paper explores the management application strategy of IoT in the key aspects of digital transmission in enterprises. It can also increase productivity to address labour shortages. Most importantly, because of the internet of things technology, many emerging industries have been derived. In order to achieve this research objective, firstly, based on the in-depth study of the problems of digital transmission of key aspects of the enterprise, this paper evaluates and summarizes the supply chain, safety, efficiency, and quality problems that often occur in the past operation and management of the case enterprise and analyses the main reasons for the formation of the problems and then recognizes the necessity and feasibility of introducing IoT technology and management; secondly, based on the kernel of IoT and industrial IoT technology combined with the main modules of enterprise operation management, using the 5S field management method, we try to build a three-dimensional all-round electrical equipment enterprise IoT management application scheme design system; and finally according to the planning and design rules and management process, we focus on implementing digital transmission network application projects from the enterprise level combining key links of technology and management. This paper studies not only technical innovation but also a revolution in management. The research in this paper will provide a more feasible technical and management design scheme for the application of IoT in enterprises for industry reference.

1. Introduction

The role played by the new generation of technology products in human life is irreplaceable, and the former science fiction movie scenarios are being used in modern life one by one [1, 2]. According to a study by the Internet Data Center (IDC), by 2020, there will be 29.5 billion connected devices worldwide, creating an IoT market of $1.7 trillion. According to a study by McKinsey & Company, a world-class consultancy, the economic impact of the IoT will be worth $11.1 trillion to $13.9 trillion per year by 2025, with industrial applications worth $1.2 trillion to $3.7 trillion, making it the largest economic application of all [3]. Many countries have included IoT as an economic development priority in their national strategies. This shows that IoT is the

future technology trend in human development, and its application not only can improve the quality of products and services and increase the source of customers but also can increase productivity to solve the labour shortage. Most importantly, many new industries have emerged because of IoT technology [4].

In recent years, the emergence of IoT technology has improved people's quality of life, and many countries have incorporated IoT into their national strategies for economic development, especially in manufacturing, which is predicted to be the largest economic application area in the future [5]. In the decade since its emergence to the present day, blockchain technology has grown rapidly into mature and practical information technology [6]. Blockchain can be thought of as a distributed ledger that is monitored and

maintained by all participants through the network, thus ensuring the authenticity and immutability of information. At the same time, blockchain, as a "decentralized trust network," can maximize the elimination of human intervention, and each node in the network does not need to know who the other is if it conforms to the consensus encryption mechanism to complete the operation and achieve distributed storage [7]. Blockchain's decentralization, immutability, openness, and transparency provide a new path for the development of digital transmission of key aspects of enterprises.

With IoT technology as the core, embedded development technology, wireless sensor network technology, and fault diagnosis technology are applied to realize the wireless monitoring and evaluation system of the pump group based on IoT [8]. The sensor is used to monitor the running state of the centrifugal pump set, and the running data of the centrifugal pump in the current state are collected and uploaded. The cloud server performs health assessment, fault diagnosis, and fault prediction on the current state of the centrifugal pump and gives corresponding treatment measures and treatment suggestions according to the system's evaluation of the current state of the centrifugal pump. This is conducive to the timely detection and troubleshooting of faults by the staff and prolongs the service life of the equipment.

2. Related Works

The industrial internet of things (IoT) is an information technology based on the Internet, which integrates various sensors, intelligent terminals, and other advanced devices by various means to achieve the interconnection of everything [9]. As shown in Figure 1, the industrial IoT architecture can be roughly divided into three layers: the sensing layer, the network layer, and the application layer. The perception layer of industrial IoT consists of many sensors and controllers, mainly including various types of sensors such as vibration, pressure, temperature and humidity, RFID tags, QR codes, and various intelligent devices. The main task of the sensing layer is to realize the collection of data from various devices and transmit the collected data to the network layer, receive and execute the commands from the network layer, realize intelligent control, and realize a wireless monitoring and evaluation system for pump groups based on the internet of things. The sensing layer of industrial IoT consists of many sensors and controllers, mainly including various types of sensors such as vibration, pressure, temperature and humidity, RFID tags, QR codes, and various smart devices [10]. The main task of the sensing layer is to realize the collection of data from various devices, transmit the collected data to the network layer, and be able to receive and execute commands from the network layer and realize intelligent control. It is mainly responsible for the transmission of various types of information and is the link for the transmission of information throughout the industrial internet of things [11]. The application layer is mainly responsible for data storage, data processing, analysis and decision-making, and so on. It controls the perception layer

to carry out corresponding actions and realize remote monitoring and control of industrial equipment to achieve the purpose of human-computer interaction in IoT.

From the perspective of the internet of things technology, its origin, development, and popularization in foreign countries started early, so the overall system technology supported by the internet of things technology applied in different directions and fields is also completer and more comprehensive. It is of great significance to give corresponding treatment measures and treatment suggestions, which will help staff find and eliminate faults in a timely manner, prolong the service life of equipment, and reduce the number of equipment downtimes. Internet of things technology is a multi-disciplinary, comprehensive technology system. It involves computers, with wireless communication, multinetwork interconnection, sensor collection, and integration technology [12]. Foreign scholars and even development teams belonging to companies have invested a lot in hardware technology support or software application and correlational and collaborative research, and some technologies are even in a forward-looking position in the world. This network is usually part of an ecosystem (ecosystem) that includes connectivity between things, applications, and data analysis. Although networks are not a prerequisite for internet systems, the value of the internet of things being realized has the potential to create a huge and global impact [13].

Concerning the impact of IoT on business operation and management, industrial IoT has indeed developed rapidly and vigorously in recent years, and production automation and process automation have improved quality and efficiency for many enterprises, in addition to which the return on investment has also been quickly obtained [14]. In this paper, we propose and study an enterprise key link digital transmission system based on the internet of things, which is supported by the three-layer structure of the internet of things as a big framework, involving node data collection, RFID access technology, data transmission module, and so on. The main components of the data are equipment status information, node water pressure at the collection end, and hardware control information; the system is intended to complete accurate data collection and smooth transmission at the same time to data information for additional processing, so that it can provide decision support for the overall system and the construction of the ontology-based monitoring and early warning system model. In terms of the data visualization user side based on the cloud data transit, a friendly interface, humanized interactive, and easy-to-operate program are developed for the web application side of the manager [15].

3. Enterprise Critical Link Digital Transmission System Model Setup

3.1. Mathematical Modelling. The overall supply chain traceability system architecture can be divided into three layers, mainly including the data access layer that obtains product data and status information and uploads them through the interface program; the data core layer that
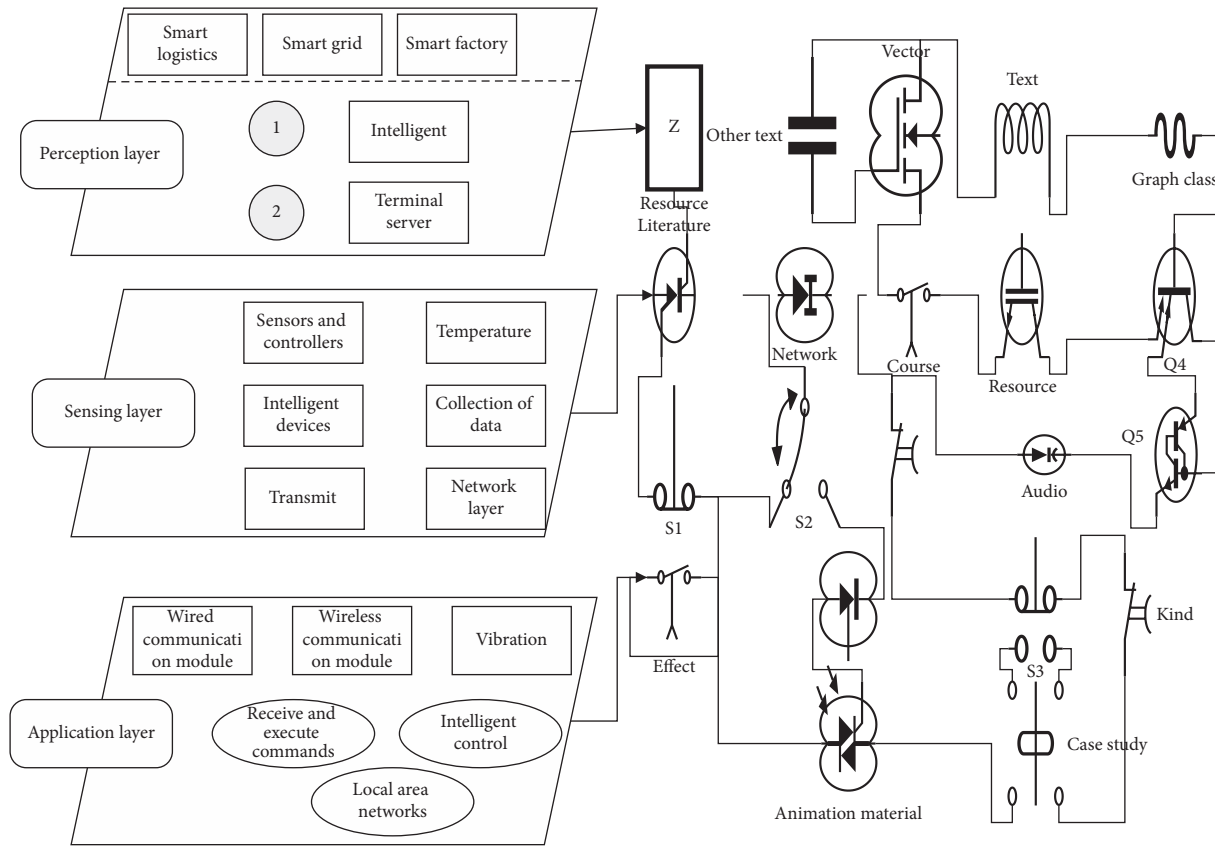
Figure 1: Simple architecture of industrial IoT.

invokes smart contracts to complete the supply chain business logic and conducts blockchain data storage according to the received data information; and the application representation layer that includes the traceability query web page and WeChat traceability query applet [16]. It involves node data collection, RFID access technology, data transmission module, and so on. The main components of the data are equipment status information, collecting end node water pressure, and hardware control information. The system aims to complete accurate data collection and smooth transmission. Information undergoes additional processing so that it can provide decision support for the overall system.

The data access layer consists of various sensors, 4G modules for data upload, and the data upload app, which mainly completes the task of acquiring product data information and uploading it to the core layer. The ways to obtain the product data information include various types of sensor acquisition and scanning the product QR code using the data upload app, which has a unique QR code number, which is the unique identity of each food product, and the QR code number is generated by 32 bit MD5 encryption algorithm; the food product is photographed using the data upload app to obtain the actual photo. The type of data information acquired mainly includes the production or growth environment information collected by sensors, the product status information acquired by scanning the QR code, and the physical photos acquired by taking photos.

The main body of the data core layer is the blockchain network, aided by a generalized interface program and smart contracts, and the entire blockchain network consists of many blockchain nodes. The full-node server that fetches the uploaded information from the data access layer calls the storage function in the smart contract through the interface program to populate the information into the transaction data and encrypt the transaction data; then the transaction is published and broadcasted in the whole network; and then the transaction is packaged into a block and broadcasted to the whole blockchain network by the consensus algorithm verification of other nodes, and the block is added to the blockchain as the latest block after the completion of the verification. Figure 2 represents the data storage process using the smart contract approach.

The user is registered through the user management module at the end of the web page and must fill in the required fields; the whole process is monitored by the administrator; the administrator account is assigned by the senior level of the system and does not have to be registered; and the new user information generated by the module will be stored to the cloud centre data [17].

Administrator users can obtain information through two-way data transmission with the database, then monitor different types and locations of node equipment, alarm equipment, and so on, and can view the status of the equipment and the adjustment of operable nodes. When a device fails, the device administrator can view the failed
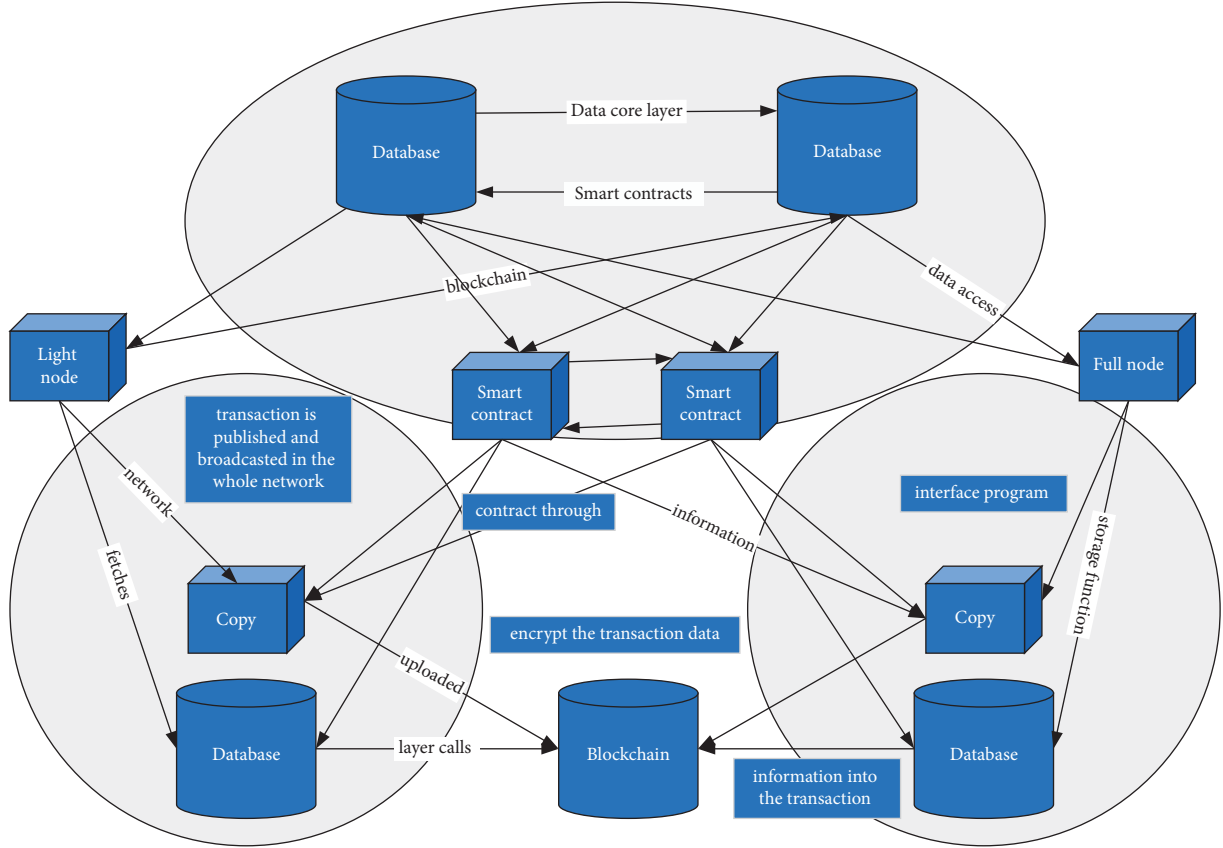
FIGURE 2: Data core layer architecture topology.

device without showing the cause of the device failure to the field and then request maintenance on the failed device. It mainly includes the data access layer that obtains product data and status information and uploads it through the interface program; the data core layer invokes smart contracts to complete the supply chain business logic, stores blockchain data according to the received data information, and includes the traceability query web page and the application presentation layer of the WeChat traceability query applet. After the approval is completed, the system will feed back the approval result. After the approval, the relevant technicians will be arranged to carry out maintenance and repair and upload the data records to the database for viewing after completion.

$$PR_T - W + C + R_D > 0. \qquad (1)$$

This paper will be based on the modern production management concept and adopt the basic idea of joint optimization of management tools and technology, based on the characteristics of motor production in the information technology environment, for the current production operation of the Shanghai Electric Machinery Factory to carry out an in-depth study of the existing problems, based on which an improvement program is proposed: The design focus of this Internet of Things application scheme is the combination of production management methods and Internet of Things technology; this theoretical framework helps promote the deeper application of traditional

production management methods and the mechanism of reuse in the introduction of IoT. The overall idea is to combine the theories of IoT and production management with the main problems of production management and operation management of Shanghai Electric Factory. The solution starts from the needs of the production operation of the enterprise, and then the architecture of the system is carried out.

Another constraint on efficiency is site management, and 5S is a powerful tool. Companies can improve efficiency in the production process while employees can work safely, healthily, and comfortably, and at the same time, companies want to reduce costs; they must start with manufacturing costs that have a high-cost structure. The product cost of the company accounts for 80% of the total cost of the company, while the manufacturing cost (direct labour and equipment improvement) accounts for about 20%, so this study will focus on the improvement of productivity through 5S to do research. In the combination with the internet of things (IoT), the main performance is that it can be supplemented with 5S management while using IoT. The full-node server that gets the information uploaded by the data access layer calls the storage function in the smart contract through the interface program to fill the information into the transaction data and encrypt the transaction data and then broadcast the transaction in the entire network, and then pass the data of other nodes. After the consensus algorithm is verified and the transaction is

packaged into blocks, it is broadcasted to the entire blockchain network. In the process of 5S management, IoT technology is actively used, such as through the completion of RFID reading and writing method tests, business process combing, and so on, to provide technical means for on-site management. Simple simulation once, the fluctuation of the adaptive alarm valve is shown in Figure 3.

Finally, through all the data collection devices on-site, the production quantity, yield rate, work-in-process distribution, equipment operation, personnel efficiency, equipment efficiency, and quality inspection results, and other real-time on-site consultation is fed back to the ERP system. Therefore, to grasp the production status of the manufacturing site, information collectors, such as handheld barcode machines or fixed barcode machines, are usually installed at each production line input station to collect various production information, operation status, environmental data, quality data, and so on through these barcode machines and then store this information in the database of the MES system. You can view device status and adjust operational nodes. When a device fails, the device administrator can view the faulty device without showing the cause of the device failure to the site and then apply for maintenance of the faulty device.

$$C = \frac{E_{\min}|x(t)|}{\sum_{t=1}^{\sqrt{n^3}} x(t^2)}. \tag{2}$$

3.2. Design of Digital Transmission Systems for Key Aspects of the Enterprise.
The cost and energy consumption of LoRa wireless communication technology are lower than those of NB-IoT wireless communication technology, and in the actual industrial application, the branch units are often located in remote places, so the NB-IoT technology may not be covered by the operator, so LoRa technology is used as the communication method for wireless monitoring in this paper [18].

$$\delta = \frac{g}{d},$$

$$d = \sum_{i,j=1}^{N} \max d_{ij}. \tag{3}$$

In MATLAB 2015a software, consider a random deployment of 50 sensor nodes in a fixed square area with a side length of 100 m. These nodes are static and do not have information about their location. A mobile NB-IoT node joins this cluster and is first at position 1 (0, 0), then moves to position 2 (33, 0), position 3 (67, 0), position 4 (100, 0), position 5 (100, 33), position 6 (100, 67), position 7 (100, 100), position 8 (67, 100), position 9 (33, 100), position 10 (0, 100), position 11 (0, 67), and position 12 (0, 33). The single- or multi-hop information from all nodes in the region is received at each location stop, and this one with the smallest value of node hops passed from each node is recorded and noted as the shortest path. The node distribution is shown in Figure 4.
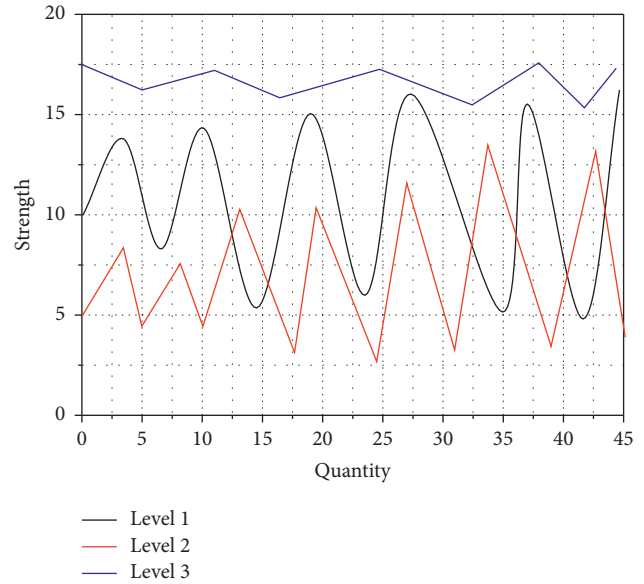
FIGURE 3: Fluctuation of the adaptive alarm valve.
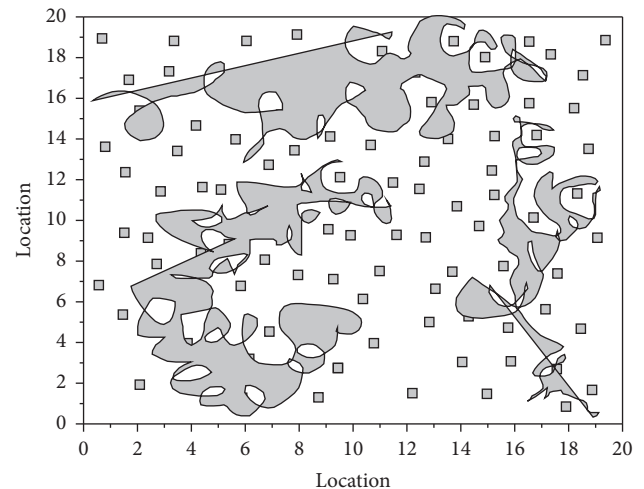
FIGURE 4: Distribution of nodes.

In Figure 4, it is the location of the mobile cluster head localization stay of NB-IoT, and ○ is the node whose location is unknown and needs to perform the localization calculation. The minimum value of the two-two combined hop count sum between any node to the 12 anchor locations is used as the shortest path between the two locations by 12 times measured hop count calculation. The application scheme of the internet of things is proposed, which is composed of multiple modules. The solution starts from the needs of the production and operation of the enterprise and then carries out the architecture of the system. An estimated jump distance value is then calculated by equation. After calculation, the node positioning error accuracy is 32.81% on average for 20 times, which achieves the position positioning function of the unknown node, but the error degree is still high. Through the analysis of the influence of related factors and parameters on the accuracy, a better combination of

positioning-related parameters is obtained for more accurate positioning.

The importance of smart contract security is increasingly recognized with the occurrence of well-known blockchain projects such as the DAO and Parity multi-signature wallets whose market capitalization evaporated overnight due to contract vulnerabilities that were attacked and caused huge losses. Formal verification is a security audit method for testing the security of smart contracts [19, 20]. Formal verification is a method that uses mathematical models to logically test hardware or software and predict results. For smart contracts, performing formal verification provides good assurance that the problems described in the previous section will not occur during the execution of the smart contract code, maximizing the security and reliability of the entire system.

4. Results and Analysis

4.1. Mathematical Model Results. In recent years, with the continuous development of Internet of Things technology, automatic sensing technology for tank level monitoring has achieved comprehensive development. After prudent research, it is believed that such aspects as an automatic level metering system, oil leakage and oil shortage real-time monitoring system, and so on can be effectively integrated with the help of intelligent internet of things platform while realizing the real-time transmission of oil tank information to enhance the overall monitoring and management of the equipment CNC bus level: (1) system advantages: each external site storage tank area has a wide range of distribution space, monitoring points, wiring complex, and other objective disadvantages and (2) real-time management: for enterprise decision-making, IoT technology allows more real-time correct management can be achieved; site managers can view real-time production progress, equipment status, material inventory status, and other information through the central monitoring platform to make appropriate decisions and arrangements to improve production efficiency. At the same time, through automated data collection, production data can be collected more completely, and through statistical analysis, abnormal quality conditions can be detected early, and product yields can be improved, and costs reduced. For workers, IoT technology replaces tedious and labour-intensive actions, such as meter reading and routine inspection—allowing more time for higher-value work, such as production data analysis, production efficiency improvement, equipment maintenance, and so on and allowing their expertise to be utilized.

$$P[X(t_n) = X(t_{n+1})]. \tag{4}$$

If $P_r + W - C + R_D - L > 0$, then $x = 1$ is the intelligent in WSNs the unique evolutionary stable state ESS of the node evolutionary game. In this communication setting, we set the value P to a constant. By increasing P_r or W or decreasing C, R_D, or L in equation, we can satisfy the condition that will force nodes to avoid choosing the no-forwarding strategy. As the value W increases (regardless of which policy the node chooses in the initial state), all nodes will eventually choose the forwarding policy, and the system will reach a steady state. Under normal circumstances, the transmission rate can only reach about 37.5 kbps, so LoRa communication technology is suitable for wireless monitoring occasions with many centrifugal pumps and a wide distribution range that do not require high communication speed.

When a node is in a busy (backlogged) communication environment and the node selects the forwarding policy too often, it wastes the node's time and effort. Therefore, in this case, it is more beneficial to choose the no-forwarding policy. We give the incentive r for nodes choosing the no-forwarding policy. R is the average probability of packet forwarding, $R = 1 - \beta$. If α or $1 - \beta$ decreases, R increases, and the positive correlation between α, $1 + \beta$, and R varies as shown in Figure 5.

There are various ways to increase the probability of packet forwarding in smart nodes in WSNs, such as increasing the bandwidth or frequency, improving hardware or transmission protocol algorithms, and so on. By increasing the R-value, $P_r + W - C + R_D - L$ tends to 0; then, nodes will increasingly choose the forwarding policy and the WSNs will reach the steady-state faster. In this case, if we increase P, it will not affect the final evolutionary steady state but will increase the speed of convergence to the evolutionary steady state. Conversely, if the value of P increases, it is difficult to satisfy the condition of Theorem 3 that $P_r + W - C + R_D - L < 0$, and then the WSNs will slowly reach the steady state. In this case, if we increase r, it will not affect the final evolutionary steady state but will reduce the speed of convergence to the evolutionary steady state. The company's product cost accounts for 80% of the company's total cost, while the manufacturing cost (direct labour and equipment improvement) accounts for about 20%, so this research will focus on improving production efficiency through 5S.

We set the parameters to satisfy the conditions of the theorem, as shown in Table 1. The experimental results are shown in Figures 4 and 5, from which we can observe the trend of the evolution curve of the node states in the smart nodes in WSNs.

There are various ways to combine NB-IoT technology with WSNs; one of the options is to use NB-IoT nodes and WSNs nodes in a hybrid network according to their characteristics, playing the characteristics of traditional WSNs nodes such as random deployment, flexible networking, and low price and NB-IoT's wireless wide-area access, low efficacy and long life, and strong penetration capability. NB-IoT and WSNs hybrid network with aggregation/cluster head nodes and common sensing nodes. Common sensing nodes are the common number of sensor nodes that are randomly distributed within the monitoring area. We do not have high requirements for the data processing capability, storage capability, and communication capability of such nodes, and they are often composed of inexpensive miniature sensors whose function is to be responsible for data collection and data transmission. Aggregation/cluster head nodes for data aggregation, which are much more capable than ordinary sensor nodes, and their impact on the
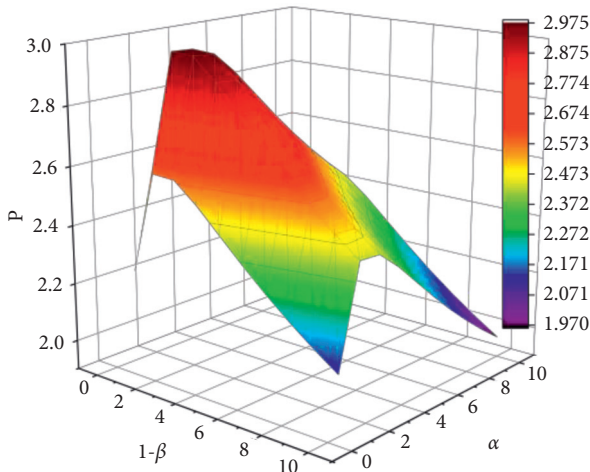
Figure 5: Plot of the variation between α, $1 + \beta$, and R.

Table 1: Experimental data.

Group	P	R_T	R_D	C	L	W
1	0.424	0.81	0.905	0.693	0.959	0.783
2	0.639	0.611	0.751	0.567	0.479	0.647
3	0.966	0.701	0.568	0.679	0.414	0.849
4	0.889	0.485	0.916	0.755	0.941	0.848
5	0.982	0.745	0.629	0.623	0.748	0.476

functional implementation of the overall WSNs is also significant. The sensor nodes send the collected data to the near-neighbour sensors, using a multi-hop transmission, and the data is finally aggregated to the aggregation node.

Wireless sensor nodes in different hierarchical positions and serving different functions have different node degrees. Sensor nodes with smaller node degrees have a single function, simple security mechanisms associated with them, and a higher probability of being controlled by viral malicious programs. Through the handheld barcode machine or fixed barcode machine, the production data, operating status, environmental data, quality data, and so on are collected, and then these data are stored in the database of the MES system. In contrast, sensor nodes with larger node degrees are fully considered in the network planning and design to be configured with system access rights restrictions, security firewalls, and other mechanisms so that their ability to cope with the prevention of attacks against cracking is enhanced, and the probability of being breached is reduced.

$$\xi(K) = \frac{m}{d(\kappa)}. \tag{5}$$

4.2. Enterprise Key Link Digital Transmission System Model Simulation Experiments. To deploy a computer as a blockchain node, we need to install node.js, the Ethereum client, the git distributed version control system, and also the Truffle compiler for smart contract writing, including Solidity, Web3.js, and Solidity development framework.

Environment variables are configured after installing Node.js, and the rest of the environment variables will be configured automatically after the software is installed. In a Linux environment, use apt-get and nm. Command in a terminal window to download and install the above software to complete the preparation work before setting up the environment.

Since EtherChannel is based on accounts to complete operations, after opening a blockchain node, the first step is to create a new account and enter personal. Click on the new account on the command line to create a new account, while EtherChannel will prompt to set a password. At this point, the account is locked and will fail when deploying and invoking smart contracts, so you need to use it personally. Since the recap set up when opening the blockchain allows access to the personal family of functions, the account can be automatically unlocked by writing this function to the Node.js interface. Before a smart contract can be called using the interface program, it needs to be compiled and deployed on the blockchain. The deployment of a smart contract on the Ethereum blockchain is done in the form of a published transaction, which is stored in the data field of the transaction. All blockchain nodes receive the transaction and perform signature verification and consensus algorithm to reach an agreement and achieve the storage of the contract on the chain. When a failure occurs in one part of the system, the system reacts and records it first, and the detailed test results of one cycle are shown in Figure 6.

By chance, we came across an AR game for preschoolers: the mobile app software scans a card with an animal pattern on it, and a three-dimensional image of the animal on the card appears on the phone with animal sounds, which helps children learn. This game uses card patterns as a medium to enhance human understanding and memory with three-dimensional images and sounds. This game has inspired us. Instead of card pattern information, QR code information or RFID information can be used in the workshop; three-dimensional images of animals can be replaced with various three-dimensional images of parts; and sound information can be presented in the form of explanations in drawings. The 2D code or RFID on-site can record various information such as the model number of the parts, their location, and their position in the finished product. The transmission distance of the signal in the open area can reach 10 km; the transmission distance in the town can reach 1~2 km; and the maximum transmission rate can reach 250 Kbps. Because there are more young workers in the workshop, most of them like to play games; the way of learning through mobile phones is easily accepted; and the three-dimensional model pattern display is far easier to be understood than explaining with drawings. Through the mobile phone can also make the content of AR learning competition software to enhance learning initiative; in addition, through the two-dimensional code mark or RFID each work step into the system, you can trace the completed work and incomplete work, to achieve the guidance function of the work task, through the mandatory process of the production process and process control, to avoid errors in the work.
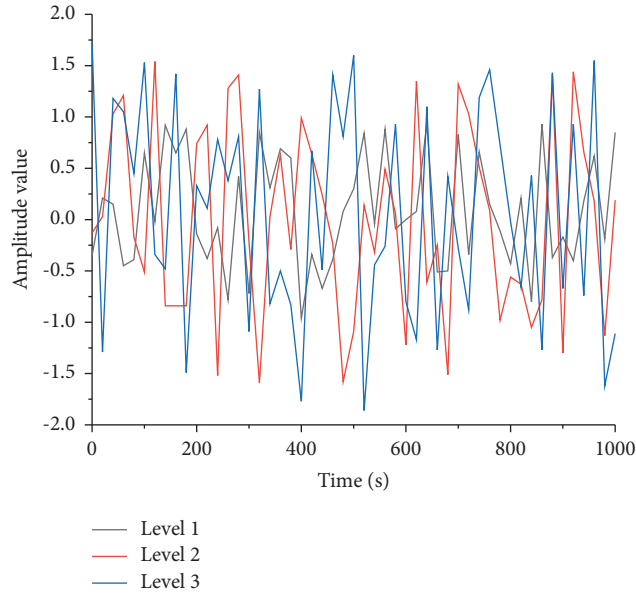
FIGURE 6: Fault time-domain data waveform.
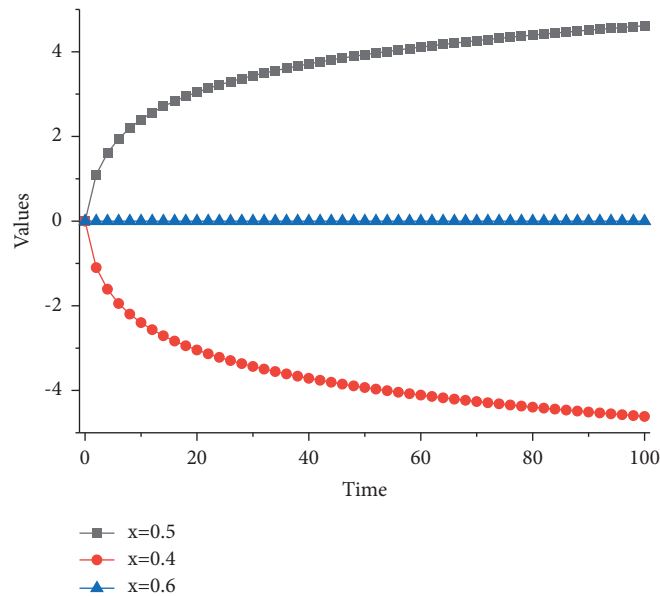
FIGURE 7: Schematic diagram of the evolution of packet forwarding curve denoted by W.

When the rate at which nodes choose a forwarding strategy in the initial state is 1%, reaching the stabilization point $\chi2^* = 1$ requires about 110 games, but if $W = 4$, reaching the same stabilization point requires only about 40 games. Satisfying the conditions of the theorem, the larger the cooperative incentive, the faster the intelligent nodes in WSNs reach the equilibrium point. When the rate of nodes choosing the forwarding strategy in the initial state is 99%, it takes more than 100 games to reach the stabilization point $\chi1^* = 0$ if $W = 3$, but only more than 80 games to reach the same stabilization point if $W = 4$. Under the condition that the theorem is satisfied, the gain of nodes that choose the forwarding strategy is smaller than the gain of nodes that do

not choose the forwarding strategy. However, the cooperation incentive promotes cooperation among nodes to forward packets to each other, thus reducing the rate at which nodes reach the evolutionary steady state of intelligent nodes in WSNs, and the dynamic trend is shown in Figure 7.

ESP-12F has three working modes: STA, AP, and STA + AP hybrid mode. From the viewpoint of the usage requirements of this system, since the monitored temperature and humidity data, voltage, and related information need to be transmitted to the cloud platform, the system must be able to connect with the router connected to the external network to realize data exchange, so the main working mode should be set in STA mode to access the Wi-

Fi network established by the hotspot already connected to the external network to realize the functions such as data upload.

The platform needs to be modularized to facilitate future expansion. Because the functions developed so far are only a very small part of the functions that the IoT can achieve and because the IoT technology is developing rapidly, it is very difficult to expand the platform later if it is not designed in a modular way, so it must be developed using a modular design. The platform must be a data presentation centre. As a platform software for industrial IoT, data presentation is the most basic function. After the data read from the sensing end are transmitted to the data collection end through the network bus, a large amount of data will be deposited in the storage; the platform needs to display two aspects of data; on the one hand, the currently collected data is displayed instantly so that employees can be informed of the current working status of the equipment; and on the other hand, historical data should also be displayed so that it can be compared with the best working status.

5. Conclusion

With the gradual launch of the 5G network, the new generation of the mobile communication network has more support and more openness for the connection of things, and the mobile communication network will play a greater role in promoting the development of IoT technology. The integration and development of IoT in various industries face a broader application prospect. Information technology, including the internet of things, is a technology for electrical equipment companies to connect customers, suppliers, and production processes; to overcome the information and communication between production equipment, so that production equipment can pass information to each other; to establish a key link between customers and suppliers through the internet of things; through the introduction of smart equipment and careful calculation of customer requirements, the production process is simplified and costs are reduced; the calculated information will be passed to production equipment and suppliers by introducing intelligent devices. We can finely calculate customer requirements; transfer the calculated information to production equipment and suppliers; achieve precise production; collect data affecting the production process; analyse, evaluate, and mine the data; and then establish a prediction model to make production equipment self-aware, self-memory, self-cognition, self-decision, and self-reconstruction. In the future, companies can use this model to produce different customized products and the lowest production cost.

Data Availability

The data used to support the findings of this study are available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] Z. Hong, H. Cai, and J. Wu, "Intelligence-oriented industrial internet of things: architecture, key technologies and challenges," *International Journal of Autonomous and Adaptive Communications Systems*, vol. 14, no. 3, pp. 264–278, 2021.

[2] W. Wu, J. Lu, and H. Zhang, "Smart factory reference architecture based on CPS fractal," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 2776–2781, 2019.

[3] K. Cebeci and Ö. Korçak, "Kurumsal ömtbmt," *Bilişim Teknolojileri Dergisi*, vol. 13, no. 4, pp. 357–371, 2020.

[4] A. Nath and S. K. Rath, "Service oriented architecture-based design of bank-ATM and its verification with Petri-net," *International Journal of Simulation and Process Modelling*, vol. 16, no. 4, pp. 343–352, 2021.

[5] I. Yaqoob, E. Ahmed, I. A. T. Hashem et al., "Internet of things architecture: recent advances, taxonomy, requirements, and open challenges," *IEEE wireless communications*, vol. 24, no. 3, pp. 10–16, 2017.

[6] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How can heterogeneous internet of things build our future: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2011–2027, 2018.

[7] T. Cerny, M. J. Donahoo, and M. Trnka, "Contextual understanding of microservice architecture," *ACM SIGAPP - Applied Computing Review*, vol. 17, no. 4, pp. 29–45, 2018.

[8] L. Wang and Y. Wang, "Supply chain financial service management system based on block chain IoT data sharing and edge computing," *Alexandria Engineering Journal*, vol. 61, no. 1, pp. 147–158, 2022.

[9] J. Liu, M. Chen, and L. Wang, "A new model of industrial internet of things with security mechanism-An application in complex workshop of diesel engine," *Proceedings of the Institution of Mechanical Engineers - Part C: Journal of Mechanical Engineering Science*, vol. 234, no. 2, pp. 564–574, 2020.

[10] Y. Lu and L. Da Xu, "Internet of Things (IoT) cybersecurity research: a review of current research topics," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2103–2115, 2018.

[11] K. Yi, J. Cao, and W. Pu, "Architecture of digital integrated R & D platform for military research institutes," *Converter*, vol. 2021, no. 7, pp. 847–858, 2021.

[12] F. A. Khan, S. Shaheen, M. Asif, A. U. Rahman, M. Imran, and S. U. Rehman, "Towards reliable and trustful personal health record systems: a case of cloud-dew architecture based provenance framework," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 3795–3808, 2019.

[13] M. Alam, J. Rufino, J. Ferreira, S. H. Ahmed, N. Shah, and Y. Chen, "Orchestration of m using d and edge computing," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 118–123, 2018.

[14] M. Nicola, C. Nicola, and M. Duta, "SCADA systems architecture based on OPC and web servers and integration of applications for industrial process control," *International Journal of Control Science and Engineering*, vol. 8, no. 1, pp. 13–21, 2018.

[15] W. Jun, "Research on the framework of smart city operating system based on new ICTs," *American Journal of Artificial Intelligence*, vol. 4, no. 1, pp. 36–41, 2020.

[16] G. Liang and W. Li, "A novel industrial control architecture based on Software-Defined Network," *Measurement and Control*, vol. 51, no. 7-8, pp. 360–367, 2018.

[17] L. A. Cruz Salazar, D. Ryashentseva, A. Lüder, and B. Vogel-Heuser, "Cyber-physical production systems architecture based on multi-agent's design pattern-comparison of selected approaches mapping four agent patterns," *International Journal of Advanced Manufacturing Technology*, vol. 105, no. 9, pp. 4005–4034, 2019.

[18] H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, and M. Alizadeh, "The application of internet of things in healthcare: a systematic literature review and classification," *Universal Access in the Information Society*, vol. 18, no. 4, pp. 837–869, 2019.

[19] F. Tao, Q. Qi, L. Wang, and A. Y. C. Nee, "Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: correlation and comparison," *Engineering*, vol. 5, no. 4, pp. 653–661, 2019.

[20] M. N. Jahantigh, A. M. Rahmani, N. J. Navimirour, and A. Rezaee, "Integration of internet of things and cloud computing: a systematic survey," *IET Communications*, vol. 14, no. 2, pp. 165–176, 2019.

Hindawi

Research Article

SDN Control Strategy and QoS Optimization Simulation Performance Based on Improved Algorithm

Bin Zhang ⓘ and Xin Liu

School of Information Engineering, Shandong Youth University of Political Science, Jinan, Shandong 250103, China

Correspondence should be addressed to Bin Zhang; 140011@sdyu.edu.cn

In the past few years, the Internet has become more and more popular, and more and more people use the Internet. A lot of new software has been developed on the Internet, so the network traffic has increased rapidly. A company once clearly stated that, in the future, the network volume in our country will reach nearly 5 zb. In addition, the network is comprehensively unified, and the learning of the Internet is strengthened by the way of suggesting errors. It is also necessary to learn from the external environment to make the finite value of the network as good as possible. Reinforcement learning is about solving many difficult problems. Use reinforcement learning to promote and change this state, so that the advantages of the network can be fully developed. Regarding the environment and actions at the time as the basic mapping requirements, and the network technology strategy can be best developed. In order to solve the QoS optimization scheme of the current mainstream heuristic algorithm in the software-defined network scene, the software-defined network QoS optimization algorithm is proposed. First, the network resources and status information are unified into the network model, and then, the long- and short-term memory network is used to improve the algorithm's flow perception ability. Finally, based on the deep reinforcement learning algorithm, a dynamic traffic scheduling strategy that satisfies the QoS objective is constructed. Among them, QoS refers to the main functions of the entire network system, especially for some new users, but for many old users, this system can represent a certain application, for example, whether a certain software can get timely, whether it can work smoothly when processing videos, and whether it can be interrupted when making a voice call for improvement.

1. Introduction

In the upcoming 2022, the total amount of China's network will reach 5zb. This number represents that the Internet is being widely used, and basically the Chinese are using the Internet from door to door [1]. It also proves that the number of people using the Internet is innumerable. Faced with the rapid development of the Internet, on the premise that many large companies and consumers use the Internet at the same time, Internet companies are also undergoing an unprecedented huge test [2]. This is because when many consumers use the same network at the same time, problems such as freezes, unclear videos, and unusable traffic will occur [3]. At present, the basic meaning of software-defined networking is to accept the major changes of the newly developed network [4]. Because the original network also

has some problems, we have to find many new ways to solve these original problems. The basic purpose of the software is to unbind the initial network control, separate it from the original, and then use the control panel to perform a comprehensive search, so that our program will not be stuck or unclear status [5]. This program was then sent to the company above, in order to unify the network in an all-round way, so as to facilitate the management of the staff and make the network develop in an orderly manner [6]. The main reason is to make network management simple because many people still do not know how to use the network [7]. If it is too difficult, they will face many problems. After solving the difficult problems, the cost must be minimized, so that our costs are reduced, and the highest benefits can be obtained, and consumers can use it with confidence [8]. We must also greatly support the advantages of network

innovation and then find out many shortcomings and solve them one by one [9]. There are many large and well-known Internet companies that are also applying this feature to give full play to their advantages, learn from each other's strengths, and maximize the interests of consumers [10]. This can also dig out potential functions on the largest scale and achieve utilization. The biggest is to change some of the functions of the network and to properly solve the previous bad aspects, so that we can use it comfortably. The goal is to make the network more meaningful, and the meaning of the network itself lies in letting consumers solve problems, make their lives more convenient, and meet the interests of consumers [11]. Only when the Internet is done can our lives become more colorful and beautiful.

2. Related Work

Literature [12] proposes that there is a lot of false information on the Internet, because the Internet itself is virtual, and many things you see may not necessarily mean that it is real, because some consumers do not understand this aspect. At the same time, many criminals have processed network information, so that many consumers have been deceived by false information on the Internet. To solve this difficult problem, the relevant departments have come up with a solution. This method is used to connect the Internet and resources. The hub of the Internet is to solve the virtual existence of the Internet, help consumers not be deceived, and strengthen online learning. Literature [13] concisely put forward some problems existing in the network and based on these existing problems came up with a solution and then constructed a model, which is called a mathematical model. Literature [14] provides a mechanism that can manage and solve this problem well. In Literature [15], it is pointed out that because the user's choice is different, each solution is different, and the strategy is also different. Literature [16] has good control ability and good calculation ability, so the content designed according to its characteristics is different, and the network path and change system are also different. Literature [17] chose to introduce a new type of technology, which can change the original resource and network allocation. In the literature [18], there is a model that proves that the use of management methods can solve the problems well. Literature [19] uses it to link with broadband in a wireless network. With broadband, half of the problem can be solved in this way. Literature [20] pointed out that there are many places where the network is used, and many places have problems that are difficult to solve. Aiming at this series of problems, an organization can solve the current problems, and its name is called the regional network. Because each piece of network is divided into different regions, each region has its own network, which solves some of them well. And because the network sometimes has disconnection problems, problems like freezing of videos occur. Literature [21] proposed software that can use algorithms to control the virtuality of the network and the allocation of resources and the network. The ultimate goal is to solve the problem of insufficient allocation. Literature [22] applies governance technology to the network and proposes two solutions. Literature [23] pointed out that the use of open data to manage

the traffic required in the network will solve the problem of unclear video network stalls. It also proposed a document mechanism to further ensure the high quality of the video and a better quality of the traffic network in the increasing number of traffic videos and the most basic difference, so as to meet the needs of consumers to a greater extent. A good network and network companies are also improving, and our lives are also progressing and will become better.

3. SDN Network Analysis Based on Deep Reinforcement Learning

3.1. SDN Control Strategy Based on Deep Reinforcement Learning. Reinforcement learning is to use an intelligent method to learn, using the environment as a guide for the next step, and its ultimate goal is to better reflect the intelligence and maximize the benefits. Through reinforcement learning, many ways to solve the problem have been discovered. One of them is because there is little information obtained through the external environment, and it must rely on its own experience and original materials to learn, and it has obtained it in the subsequent environment. Reinforcement learning can obtain a lot of information from the environment, so as to analyze and change the current environment. There is an algorithm that can be used as a major point of solving the problem, and that is to calculate the resource allocation in the wireless network. This calculation should be solved first. In the algorithm, we have to make analogies and assumptions and then calculate the number of calculations. The results are calculated, and the $\triangle$ is the most important core. Observe and get rewards or punishments. The following are rewards or punishments:

$$R_t = \sum_{i=t}^{T} \gamma^{(t'-t)} \cdot \gamma_i. \tag{1}$$

Among them, y belongs to a factor, and R is calculated from beginning to end and includes the sum of rewards and the sum of punishments. In addition, t means time and q is defined as a special dynamic variable in mathematics. The core of an algorithm is to use formulas for algebra. Use this method to learn and use simpler methods to calculate dynamic variables:

$$\begin{cases} Q_{k+1}s_t, \alpha_t = Q_k s_t, \alpha_t + \alpha_k \cdot \delta_k, \\ \delta_k = \gamma_{t+1} + \gamma \cdot \max Q_k s_{t+1}, a' - Q_k s_t, \alpha_t \, (a' \in A). \end{cases} \tag{2}$$

In the formula, α represents the time and rate of learning. In addition, s and a, respectively, represent the distance and a certain state of the action, which is the difference in time, and α represents the state and distance of the execution n times. In addition, in order to realize the algorithm of formula two, the conditions must be met:

$$\begin{cases} \sum_k a_k^2 < +\infty \, , \quad \sum_k \alpha_k = +\infty, \\ \lim_{k \to \infty} Q_k = Q^*. \end{cases} \tag{3}$$

According to the process described above, we can know that in a certain state, and we can choose the best method by ourselves; that is, choose the best expected value to analyze a specific function:

$$Q^* s, a = E_{s'-\xi} (r + \gamma \cdot max) Q^* (s', a'|s, a). \tag{4}$$

In the general environment, various states will exist. s is regarded as a certain state after performing an action, and α is a combination of one state and another state and executed.

3.1.1. Constructing Special Value Functions.
By strengthening the learning of the special value function in the network, taking this function as the core problem in transforming it into another function, it can be expressed as

$$\hat{Q} s_{sys}, a; \Theta \approx Q^\pi (s_{sys}, a). \tag{5}$$

In a more in-depth research algorithm, if you want to calculate a certain function, you can rely on a model. This model can update itself and then calculate it. The special value function in the Q-Learning algorithm can be rewritten as

$$Q^* s_{sys}, a = Q s_{sys}, a + a \left[R_{once} + \gamma max Q s_{sys}, a' - Q(s_{sys}, a) \right]. \tag{6}$$

Among them, a is a learning factor, and a is a learning factor in another state and a certain action in another state.

A certain kind of table cannot store every state of the system, so we have to change this kind of table and solve the solution in this article, which is to let the functions replace each other and express them by approximate functions.

3.1.2. Constructing Loss Function under Dual Network Architecture.
To make the results calculated by this algorithm more specific, two very similar network mechanisms must be constructed. One of the mechanisms is relatively stable and does not change frequently. We call it the destination network and use a letter to value it. Another mechanism is called the core network. Determine a certain function by strengthening the computing power, and its loss function expression is

$$\lim_{k \longrightarrow \infty} Q_k = \Theta_*. \tag{7}$$

A function gradient can be obtained by calculating through listing:

$$\frac{dL\Theta}{d\Theta} = E \left[Q_{target} - (s_{sys}, a; \Theta) \frac{dQ(s_{sys}, a; \Theta)}{d\theta} \right]. \tag{8}$$

Use the previously mentioned formula to calculate the function gradient for a lot of training and finally get the most accurate function value, and the purpose is to achieve the goal of special points. There is a special derivative:

$$\nabla_{\Theta_k} L_k \Theta_k = E \left[\left(R_{once} + \gamma max Q_{main} s_{sys}', a' ; \Theta_{k-1} - Q_{main} s_{sys}, a; \Theta_k \right) \nabla_{\Theta_k} Q s_{sys}, a; \Theta_k \right]. \tag{9}$$

Among them, the value of one parameter is fixed.

3.1.3. Experience Playback.

$$Q s_{sys}, a; \Theta_k \overset{modification}{\longrightarrow} Q(\psi s_{sys}, a, \theta'). \tag{10}$$

In addition, an input state in the ΦBei network system is also a combination of two parameters including neural network and destination network. The corresponding mechanism can be modified to

$$\begin{cases} D \overset{modification}{\longrightarrow} \overline{D} = [\overline{e_1}, \overline{e_2}, \dots \overline{e_l}], \\ \overline{e_i} = (\psi s_{sys,i}, a_{net,i}, r_{once,i}, \psi s_{sys,i+1}). \end{cases} \tag{11}$$

3.2. Analysis of SDN Framework Modules

3.2.1. QoS Routing. Topology Management Module.
There is a module that uses an automatic check for network problems to identify the connection between the controller and the noncontroller switch and the switch and automatically write down the Internet address, IP address, and open data and other information, and check information: save, delete, update and change regularly.

Network Monitoring Module. One reaction in the network is to block the network traffic. The more serious the blockage is, the stall will occur. Use (12) to calculate the link congestion rate:

$$g_{link} = \frac{C}{B}. \tag{12}$$

In Table 1, B represents the amount of stutter in the video, C represents the amount of bandwidth in the link, and g represents the sum between the two quantities. Therefore, if we want to calculate this number, we must find the calculation number in it. The China map finds that it can be represented by the amount of special points and the amount of broadband at special points. The acquisition of the open protocol is simple, and it can be obtained by receiving and sending through the port. If you want to obtain the flow table, you must also use the calculation shown in Table 1.

When performing statistics, you can only select the port and number of special points, the number of bytes, and the time included. At t_1, the transmission of a special point is represented by a letter, and the number of bytes received is R. At time t_2, the transmission of a special point is represented by another letter, which represents a received time, and finally find the corresponding result. The finally

structofp_port_stats{	
uint16_tport_no;	/ ∗ The port number ∗ /
uint64_trx_packets;	/ ∗ Number of received messages ∗ /
uint64_trx_packets;	/ ∗ Number of sent messages ∗ /
uint64_trx_bytes;	/ ∗ Number of bytes received ∗ /
uint64_trx_bytes;	/ ∗ Send bytes ∗ /
uint64_trx_bytes;	/ ∗ Duration ∗ /
};	

obtained network monitoring data can be expressed by Equation (16):

$$B = \frac{Tx_2 - Tx_1}{t_2 - t_1} + \frac{Rx_2 - Rx_1}{t_2 - t_1}. \tag{13}$$

We can also calculate the result we want using formulas (15) and (16).

The open protocol does not have a way to directly detect the network delay, so it is not a good method to use the open protocol to calculate the delay.

Path Calculation Module. Use the methods passed through to calculate the comprehensive network information and use the data provided by the network monitoring module. Relatively speaking, the highest-level data is used for statistics, and other low-level data should be used with the least. The path is calculated at the cost.

Routing Management Module. By controlling the router to calculate the number we want and send it to the company for the company to carry out the next exchange, the switch can carry out the next business forwarding. It mainly contains the following functions:

This module can identify different businesses differently. After identifying the highest and best data, it will be sent to other companies and the results will be calculated. Compared with other levels of data, the results will be sent out. The best result can be calculated at the cost of the shortest path.

3.2.2. Queue Scheduling Management

Queue Agent Module. In a modular display controller, if we want to open the company to exchange and communicate commands, this is what we want.

Traffic Management Module. The flow module is used for management in the open switch, so that the rapid channel and the template in the controller can be integrated with each other. The related queue scheduling and simple algorithm become very simple, and the management strategy can be analyzed and answered later.

3.3. Analysis and Design of Lagrangian Relaxation Routing Algorithm

3.3.1. Delay-Constrained Minimum Cost Path Problem.
There are definitely the most advanced and advanced ones in the network, and there are also some problems with delay

and congestion in the network. If you want to find this value, you want to minimize link congestion. Then, we have to figure out a way to solve it.

$$cp = \sum_{u,v \in p} c_{uv}$$
$$dp = \sum_{u,v \in p} d_{uv} \tag{14}$$
$$p^* = \min\left\{cp|dp \le \Delta_{delay} , p \in P_{st}\right\}.$$

3.3.2. Using Lagrange's Median Theorem for Algebraic Algorithms.
In real life, there are good ways to solve this problem, but it is impossible to find a theory to replace this algorithm. It is necessary to use Lagrange's median value theorem to approximate this value, so that this problem can be solved.

LARAC is implemented by an aggregate cost function, which is shown in formulas (15) and (16).

$$c_{uv}\lambda = c_{uv} + \lambda \times d_{uv}, \tag{15}$$

$$c_\lambda p = \sum_{u,v \in p} c_{uv}\lambda. \tag{16}$$

Use a letter to represent Lagrange's median theorem, and c represents a level of aggregate cost function.

Use Lagrange's median theorem to calculate link congestion rate and weighted rate; see equation (18).

$$c_{uv} = ag_{uv} + 1 - a \times 1, 0 \le a \le 1 , \forall u, v \in E, \tag{17}$$

$$\lambda = \frac{c(p_c) - c(p_d)}{d(p_d) - d(p_c)}. \tag{18}$$

If d(pr)≤Δdelay, then a new way must be found; otherwise, the solution cannot be solved by using that way.

3.4. PDW Queue Scheduling Management Algorithm.
According to a technique for randomly detecting queues, this technique can be divided into two algorithms: one algorithm is to calculate the average queue length, and the other algorithm is to calculate the probability. The detailed calculation formula is shown as follows:

$$q_{ave}t_n = wq_{ave}t_{n-1} + (1 - w)qt_n. \tag{19}$$

The probability of packet discarding is related to the minimum value min, maximum value max, and the reference probability p. The calculation formula for the discarding probability p is shown as follows:

$$p = p_b \times \frac{q_{ave}t_n - q_{th\_min})}{q_{th\_max} - q_{th\_min}}. \tag{20}$$

This algorithm is the biggest change to the previous algorithm, and it can judge the level of scheduling management and carry out the congestion processing rate. And

its change is to add a level-related processing on the original basis. And through this kind of relevant processing, it is very good to improve the service quality of the network.

4. Software-Defined Network QoS Optimization and Implementation

4.1. QoS Control Framework and Route Optimization. The router involves four modules: network management module, network monitoring module, network path calculation module, and routing management module, as shown in Figure 1.

4.1.1. Topology Management Module. As soon as the open protocol enters a certain network mechanism, the controller will respond quickly and discover the protocol and periodically control the machine and switch to exchange information with the switch and then change and update these links. The last step is that the exchange events of these links are recorded in the list.

This kind of controller uses the shortest path for calculation, and its disadvantage is that it cannot make advanced distinctions based on the shape of the network. It can also protect all kinds of information on the network from being infringed. It can also use a certain method to exchange and protect the mapping information of the node based on the comprehensive mapping information provided by the network monitoring module. The purpose of this method is to map the switch node to other switch nodes. This method is also the realization of the algorithm. After the switches are exchanged, two methods can be obtained to obtain the minimum path between the switches.

This kind of management system will make periodic changes, change the information according to this periodic change, and then perform different corresponding processing based on this information. The purpose is to protect the stability of this network. This kind of management system also has very important components such as top management. With this component, the shortest path and minimum cost between special points can be calculated. The relationship between related classes and interfaces of topology management module is shown in Figure 2.

4.1.2. Network Monitoring Module. The network monitoring module can obtain comprehensive data and network congestion rate periodically and orderly and calculate the path to maximize the processing of this information. Among the other two modules, the two network monitoring modules are also a good choice to calculate data in an orderly manner.

Module management should be performed on the link, and the controller should be used to expand the module, and the control module should be used as the most basic management module. Let other modules forward and discard the processing separately. The most important module is processing. After processing, it can be forwarded according to the path. Table 2 contains the main methods.

4.2. Experimental Network Test Platform

4.2.1. Experimental Network. The software and hardware configuration of each node in the network is shown in Table 3. In the table, DPID/IP represents the switch network identification number or the terminal host IP address.

4.2.2. Business Priority Mark. To distinguish different services differently, it is necessary to carry out a special mark for each type of service. From the following table, we can know that there are many special values for the business priority level of the business type, as shown in Table 4.

4.2.3. Experimental Testing Tools. Iperf is a network performance testing tool that can test the network data reception rate and packet loss rate. Use this software to send and receive data, and test the data receiving rate and packet loss rate.

4.3. QoS Optimization Function Test and Result Analysis

4.3.1. Experiment 1: QoS Path Calculation Validity Test

   (i) Experimental purpose
   (ii) Under the premise of network congestion, the calculation method is used to find the result with a small path cost.
   (iii) Experimental scenarios and steps
   (iv) Use the HTB queue in LinuxTC to limit the maximum sending rate of the switch port to 5 Mbps. Since the network mainly transmits UDP packets, the transmission volume of other types of packets is small, so the maximum transmission rate of the switch port is used as the bandwidth of the link connecting the port.
   (v) Experimental results and analysis

According to Table 5, we can know that if the congestion rate is different, its result will be different.

From Table 5, we can see how fast the router is, its congestion rate, its cost, and some important information. Its supply chain is shown in Figure 3. It can be seen that the data in parentheses can reflect the price of the cost, which can be obtained by calculation. So this router is composed of OVS4 and OVS1, so that the cost can be lower.

4.3.2. Experiment 2: Dynamic Rerouting Effectiveness Test

   ① Experimental purpose
   (i) When the network QoS path is congested, the path can no longer meet the QoS requirements of the highest priority service data flow. It is necessary to test whether the highest priority service data flow can be rerouted to the new QoS path when the QoS path is congested.
   ② Experimental scenarios and steps
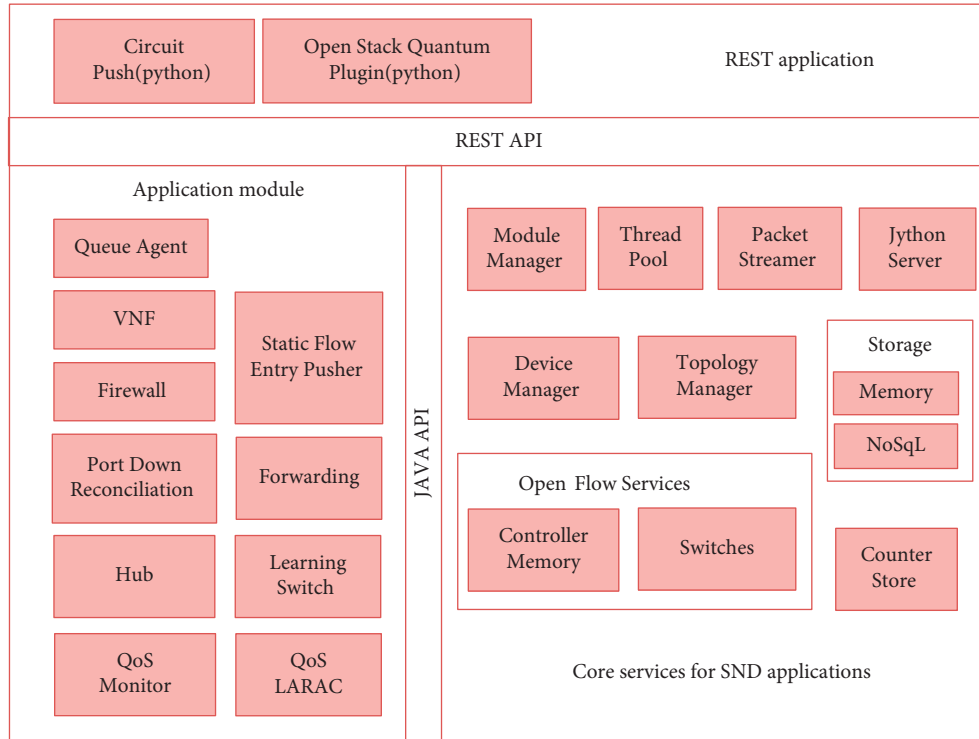
FIGURE 1: Prioritized QoS control is implemented in Floodlight.
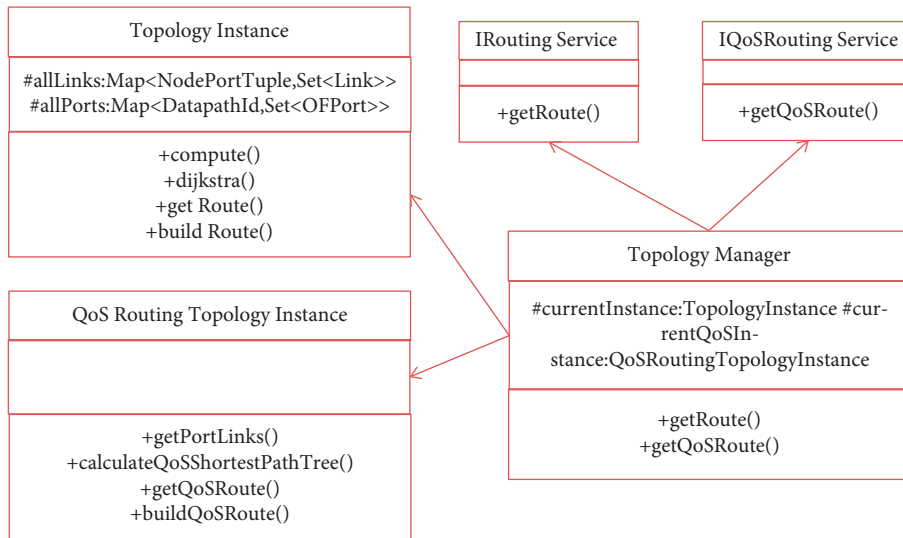
FIGURE 2: The relationship between related classes and interfaces of topology management module.

TABLE 2: Main methods of DynamicQoSRoute class.

Method	
addPath	The source switch destination switch pair is used as an identifier, a new QoS path is added to the path set cache, and the increased path time is stored at the same time.
updatePath	Update the QoS path between the specific source switch and destination switch pair in the path cache.
delPath	Delete the timeout invalid path in the path set cache.

TABLE 3: Network node software and hardware configuration.

Node	Hardware	Software system	DPID/IP
HOST1			10.0.3.2/24
HOST2	Dell-OPTIPLEX-360	Ubuntu-12.04(linux-3.13.1)	10.0.3.4/24
HOST3			10.0.3.5/24
HOST4			10.0.3.6/24
HOST5			10.0.3.22/24
HOST6			10.0.3.44/24
OVS1			00 : 00 : 00:0e:c6:c1:56:dc
OVS2		Ubuntu-12.04(linux-3.13.1) OpenvSwtich-2.3.0	00 : 01 : 02 : 03 : 04 : 05 : 06 : 07
OVS3			00 : 00 : 00:0e:c6:cb:3b:5 b
OVS4			00 : 00 : 00:0e:c6:c1:f4:1c
F1	Dell-OPTIPLEX-9020	Ubuntu-12.04(linux-3.13.1) Floodlight1.2	—

TABLE 4: Business priority mapping.

Business type	Business priority value	Transport layer source port number	Ds domain value	Tos field value
Conversational business	4	25554	4	16
Streaming business	3	25553	3	12
Interactive business	2	25552	2	8
Background business	1	25551	1	4

TABLE 5: Network link congestion rate and delay information.

Link	Congestion rate	Cost price	Time delay (ms)
E1(OVS3- > OVS1)	0.1727	26	20
E2(OVS2- > OVS1)	0.4096	46	60
E3(OVS4- > OVS1)	0.8259	84	145
E4(OVS2- > OVS3)	0.3254	39	28
E5(OVS4- > OVS2)	0.1067	19	15
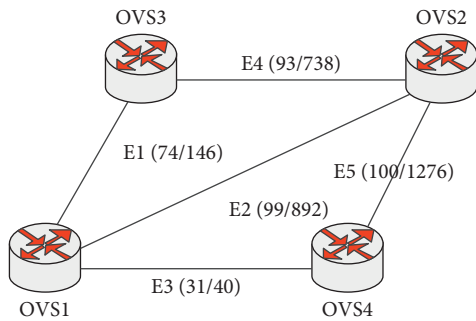
FIGURE 3: QoS routing topology.

(ii) In experiment 1, the QoS transmission path is OVS4- > OVS2- > OVS3- > OVS1, and the Floodlight controller will store this QoS path in the path set cache. After experiment one, continue with this experiment. The experimental steps are as follows:

(1) On the switch side, use the flow table delete command to delete the flow table entries in the switch.

(2) Add static flow entries in OVS4, OVS2, and OVS1, so that HOST5 sends background UDP packets with the transport layer source port number of 25551 to HOST6 through these three switches in turn.

(3) HOST5 sends UDP messages with source port numbers of 25554 and 25551 to HOST6, and the messages are sent for 30 minutes. Since the Floodlight controller path set caches the session message QoS path, the session message transmission path is OVS4- > OVS2- > OVS3- > OVS1, and the background message transmission path is OVS4- > OVS2- > OVS1.

(4) Read Floodlight log, obtain network link parameters and update QoS path.

③ Experimental results and analysis

Looking at Table 3, we can see that Floodlight calculates the QoS path as OVS4- > OVS1.

From the data in Table 6, it can reflect some accidents, costs, and important information encountered when the router is in use. When we are using the LARAC algorithm, we can calculate the final result. The information we get from now is the speed of the router, how much it costs, and how long it takes to get the network link without crowding it. It has the best service, that is, it can pass QoS, so that we can know whether the path will be updated by comparing the old and new data.

4.4. QoS Optimization Performance Test and Result Analysis

4.4.1. Experiment 3: Dispatching and Managing the Network and Inspecting Its Performance.

The higher the level, the faster the speed; the QoS performance will be best protected, which can explain the importance of performance and speed, as shown in Figures 4-5.

TABLE 6: Update the link congestion rate and delay information before the QoS path.

Link	Congestion rate	Cost price	Time delay (ms)
E1(OVS3->OVS1)	0.7145	74	146
E2(OVS2->OVS1)	0.9925	99	892
E3(OVS4->OVS1)	0.2317	31	40
E4(OVS2->OVS3)	0.9252	93	738
E5(OVS4->OVS2)	0.9999	100	1276
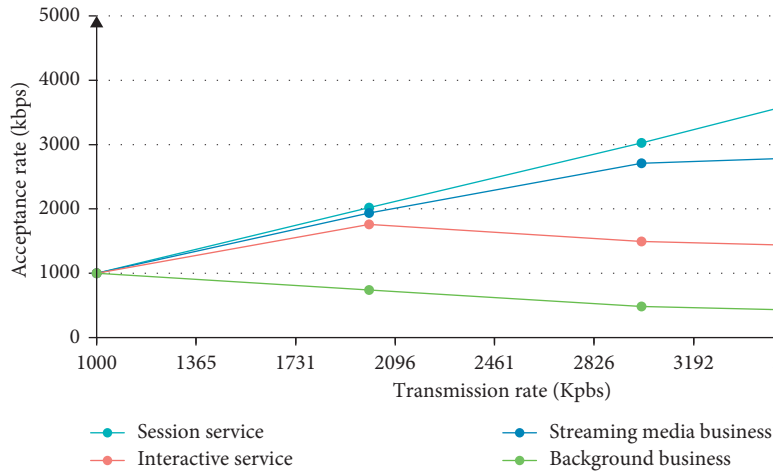
FIGURE 4: PDW policy classification service receiving rate.
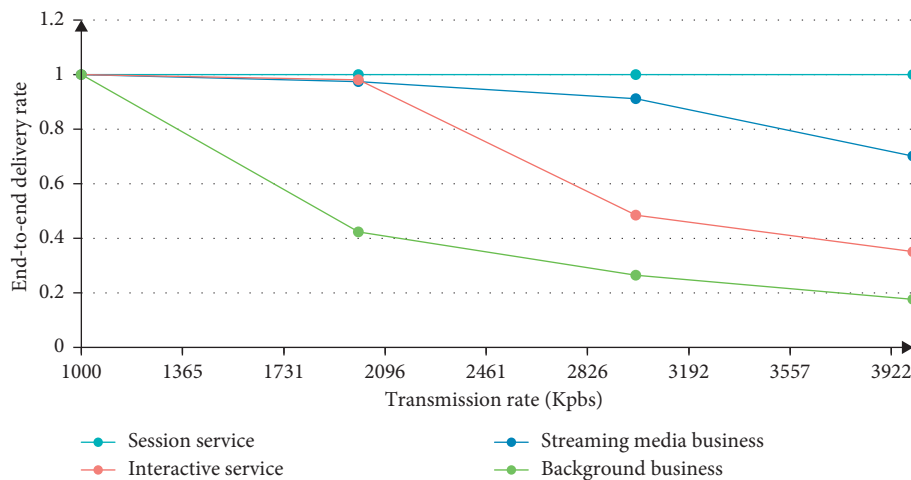
FIGURE 5: PDW strategy classification business submission rate.

The higher the level, the better the effect, and the performance of QoS can be protected in advance. This explains the requirement that the PDW policy distinguishes performance and protects different levels of performance.

4.4.2. Experiment 5: Control Frame Level and Performance Test

(i) ① Experimental purpose

In this experiment, we can obtain important information about the speed of reception and transmission. Only after this operation can we detect the benefits of framework performance.

② Experimental scenarios and steps

We can understand from the figure, the importance of broadband transmission speed. The broadband performance is good, the speed is fast, the control frame level is higher, and at the same time, a high-performance processor is required to operate, so that there will be no delay.

When we look at Figures 6 and 7, we can see at a glance the usage of the sending rate and the importance of the framework.
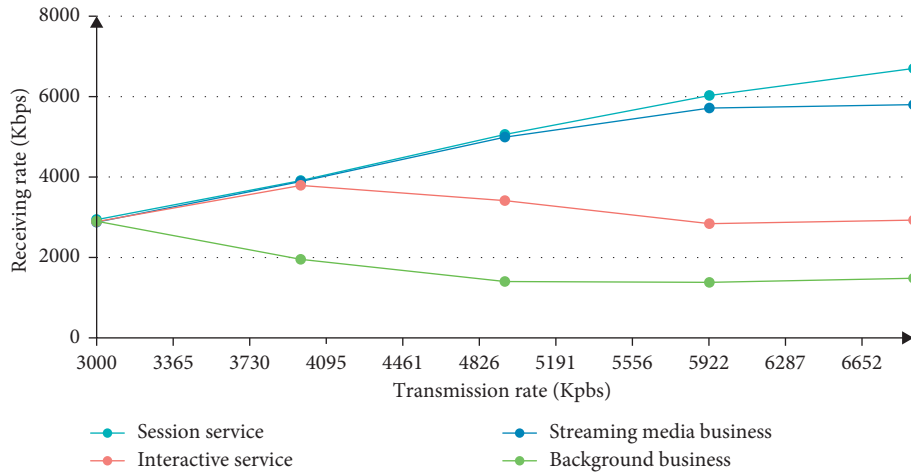
FIGURE 6: QoS control framework receiving speed of different priority packets.
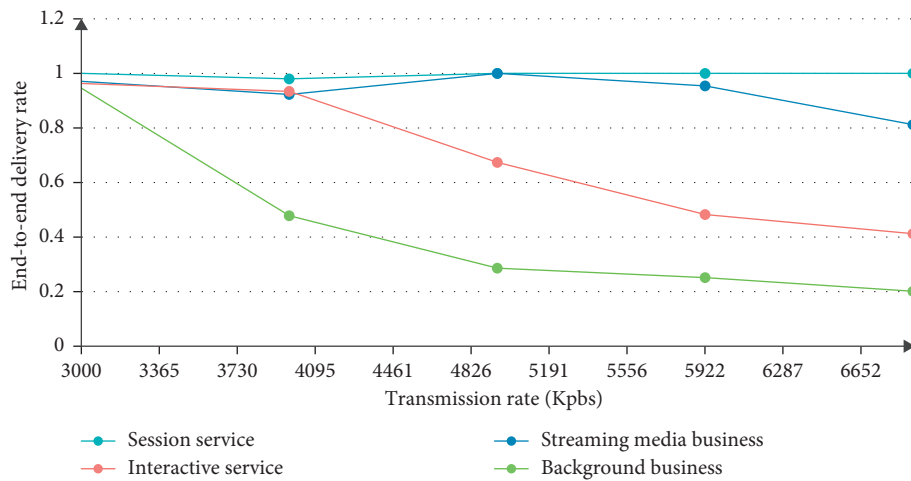
FIGURE 7: QoS control framework different priority message delivery rate.

Figure 7 shows how fast the best sending rate is, and the most advanced receiving speed can better stabilize this program. When the sending rate reaches 2 times, the data transfer speed will be very fast, and the program can be stabilized without the best telegram. Then, its stability needs a best suit to complete, making it more advanced. From this point of view, the higher the QoS framework, the better, and the QoS control framework also achieves better stability, improves network utilization, increases the amount of data transmission, and improves usage efficiency, so it can be used better.

5. Conclusion

We all know that, in today's networks, the speed of the network is fast or slow; some are very fast, and some are very slow, and even disconnected and stuck. And the quality is good or bad. Some networks are of good quality and very efficient, and some of them are not satisfactory, so we need to strengthen network learning. The speed of the network speed and the quality of the network are important factors that

affect the network, and the most advanced transmission and reception are guaranteed, and other services have received the best service from then on. This is the network we need most. In the network, we also need to study the priority levels of different services. For the most advanced clients, we must ensure timely solutions. For some low-level services, we must perform module partition processing. Research the QoS of different priority services in the SDN network, provide end-to-end delay guarantee for the highest priority services, and provide differentiated QoS services for other services. By introducing the concept of QoS, two traditional network models that can improve QoS are explained, and the characteristics of the models are analyzed. The four types of services transmitted in the network are differentiated and set according to their QoS requirements. In order to ensure the QoS of different services, a QoS control framework in the SDN network is proposed. The various functional modules in the QoS control framework are implemented in Flood-light and OpenvSwitch soft switches. From the previously mentioned situation, the QoS framework has important functions and advantages. It can make the speed faster, the

quality better, and it also promotes better service. It has been welcomed by the public, so the development of routers is a good prospect. There will be better development in the future, which can help people observe and receive information anytime and anywhere, but also the utilization rate has been well developed, without polluting the environment, and expanding the transmission volume. At the same time, it is also necessary to know how to develop a better framework and exchange tests based on this original QoS control framework, as well as the control of the control framework by the switches in the network. The switch plays a very important role in the network because it can make the best use of the various functions in the framework and finally become a reality, so that some false things in the network no longer exist and the network can function as a router. Performing a perfect test to maximize the benefits of the network is the best result we need.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *Journal of Internet Services and Applications*, vol. 1, no. 1, pp. 7–18, 2010.

[2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[3] P. Lai, Q. He, G. Cui et al., "Edge user allocation with dynamic quality of service," *Service-Oriented Computing* in *Proceedings of the International Conference on Service-Oriented Computing*, vol. 11895, pp. 86–101, Springer, Toulouse, France, October 2019.

[4] A. Blenk, A. Basta, M. Reisslein, and W. Kellerer, "Survey on network virtualization hypervisors for software defined networking," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 655–685, 2016.

[5] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (SDONs): a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2738–2786, 2016.

[6] R. Vilalta, V. López, and A. Mayoral, "The need for a Control Orchestration Protocol in research projects on optical networking," in *Proceedings of the European Conference on Networks and Communications, EuCNC*, pp. 340–344, Paris, France, July, 2015.

[7] N. M. M. K. Chowdhury and R. Boutaba, "A survey of network virtualization," *Computer Networks*, vol. 54, no. 5, pp. 862–876, 2010.

[8] K. Lam and N. Davis, "Open networking foundation, technical recommendations ONF TR-512," *Core information model (coremodel)*, vol. 30, 2015.

[9] P. Iovanna, F. Cavaliere, F. Testa et al., "Future proof optical network infrastructure for 5G transport," *Journal of Optical Communications and Networking*, vol. 8, no. 12, p. B80, 2016.

[10] J. Schönwälder, M. Björklund, and P. Shafer, "Network configuration management using NETCONF and YANG," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 166–173, 2010.

[11] 5G. Initiative, *5g white Paper," NGMN Alliance, Final Deliverable 1.0*, Next Generation Mobile Networks Ltd, Germany, 2015.

[12] M. D. Molina, S. S. Sundar, T. Le, and D. Lee, "Fake news is not simply false information: a concept explication and taxonomy of online content," *American Behavioral Scientist*, vol. 65, no. 2, pp. 180–212, 2021.

[13] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.

[14] K. E. Jeon, J. She, P. Soonsawad, and P. C. Ng, "BLE beacons for internet of things applications: survey, challenges, and opportunities," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 811–828, 2018.

[15] W. Wei, Q. Ke, J. Nowak, M. Korytkowski, R. Scherer, and M. Woźniak, "Accurate and fast URL phishing detector: a convolutional neural network approach," *Computer Networks*, vol. 178, Article ID 107275, 2020.

[16] H. A. Santoso, E. H. Rachmawanto, A. Nugraha, A. A. Nugroho, D. Rosal Ignatius Moses Setiadi, and R. S. Basuki, "Hoax classification and sentiment analysis of Indonesian news using Naive Bayes optimization," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 799–806, 2020.

[17] S. N. Mahapatra, B. K. Singh, and V. Kumar, "A survey on secure transmission in internet of things: taxonomy, recent techniques, research requirements, and challenges," *Arabian Journal for Science and Engineering*, vol. 45, no. 8, pp. 6211–6240, 2020.

[18] S. AlZu'bi, B. Hawashin, and M. Mujahed, "An efficient employment of internet of multimedia things in smart and future agriculture," *Multimedia Tools and Applications*, vol. 78, no. 20, Article ID 29581, 2019.

[19] E. Nica, K. Janošková, and M. Kovacova, "Smart connected sensors, industrial big data, and real-time process monitoring in cyber-physical system-based manufacturing," *Journal of Self-Governance and Management Economics*, vol. 8, no. 4, pp. 29–38, 2020.

[20] S. P. Sankar, T. D. Subash, N. Vishwanath, and D. E. Geroge, "Security improvement in block chain technique enabled peer to peer network for beyond 5G and internet of things," *Peer-to-Peer Networking and Applications*, vol. 14, no. 1, pp. 392–402, 2021.

[21] J. Li, A. Maiti, M. Springer, and T. Gray, "Blockchain for supply chain quality management: challenges and opportunities in context of open manufacturing and industrial internet of things," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 12, pp. 1321–1355, 2020.

[22] Y. R. Shrestha, S. M. Ben-Menahem, and G. Von Krogh, "Organizational decision-making structures in the age of artificial intelligence," *California Management Review*, vol. 61, no. 4, pp. 66–83, 2019.

[23] E. Keane, K. Zvarikova, and Z. Rowland, "Cognitive automation, big data-driven manufacturing, and sustainable industrial value creation in internet of things-based real-time production logistics," *Economics, Management, and Financial Markets*, vol. 15, no. 4, pp. 39–48, 2020.

Research Article

Towards Energy-Efficient and Delay-Optimized Opportunistic Routing in Underwater Acoustic Sensor Networks for IoUT Platforms: An Overview and New Suggestions

Varun G. Menon [1,2] **Divya Midhunchakkaravarthy,** [3] **Aaromal Sujith,** [2] **Sonali John,** [2] **Xingwang Li,** [4] **and Mohammad R. Khosravi** [5]

[1]*Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya 47301, Malaysia*
[2]*Department of Computer Science and Engineering, SCMS School of Engineering and Technology, Ernakulam 683576, India*
[3]*Center of Postgraduate Studies, Lincoln University College, Petaling Jaya 47301, Malaysia*
[4]*School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China*
[5]*Department of Computer Engineering, Persian Gulf University, Bushehr, Iran*

Correspondence should be addressed to Mohammad R. Khosravi; m.r.khosravi.taut@gmail.com

In underwater acoustic sensor networks (UASNs), the reliable transfer of data from the source nodes located underwater to the destination nodes at the surface through the network of intermediate nodes is a significant challenge due to various unique characteristics of UASN such as continuous mobility of sensor nodes, increased propagation delay, restriction in energy, and heightened interference. Recently, the location-based opportunistic routing protocols seem to show potential by providing commendable quality of service (QoS) in the underwater environment. This study initially reviews all the latest location-based opportunistic routing protocols proposed for UASNs and discusses its possible limitations and challenges. Most of the existing works focus either on improving the QoS or on energy efficiency, and the few hybrid protocols that focus on both parameters are too complex with increased overhead and lack techniques to overcome communication voids. Further, this study proposes and discusses an easy-to-implement energy-efficient location-based opportunistic routing protocol (EELORP) that can work efficiently for various applications of UASN-assisted Internet of Underwater Things (IoUTs) platforms with reduced delay. We simulate the protocol in Aqua-Sim, and the results obtained show better performance than existing protocols in terms of QoS and energy efficiency.

1. Introduction

The genesis of life on Earth had its inception on water, from which it went on to conquer varied frontiers. With the advent of the latest technologies, today's world is more connected than ever before, but ironically the blue planet still lacks efficient underwater connectivity. Underwater acoustic sensor networks (UASNs) [1–3] made their way into the limelight of research quite recently; its pivot objectives deal with an array of versatile interests from oceanographic studies dealing with marine geology, marine ecology, and physical and chemical oceanography. Another significant application is resource extraction, which mainly concerns harnessing abundant rare-earth minerals, petroleum, and natural gas under the sea bed, calamity prevention, deep-sea climate monitoring, and protection, surveillance, and reconnaissance of strategic waters by naval forces around the globe. Conventional methods used for undertaking these tasks mentioned above require humans to physically dive into the ocean's depths or rely on remotely operated underwater vehicles (ROUVs). After the data collection process gets over, these units resurface to provide

the output. The data acquired always fell short of fulfilling its objectives as there were problems like the lack of accurate real-time data, stringent storage constraints, inability in handling mobility, and capability to withstand underwater pressure.

UASNs have cell-powered sensor nodes deployed throughout the ocean bed that interact with each other and with the sonobuoys located at the surface to suffice these objectives. Their presence ensures effective communication with the sensor nodes (real time), and they are also the first responders to notify the base station if any nodes fail. Besides all these conspicuous merits, they have been used in underwater acoustic research and antisubmarine warfare for a long time, reflecting its practicality as the UASN's function under tight frequency limitations. Numerous unique features of the underwater environment make the deployment and use of UASN quite a challenging task [4–7]. Underwater conditions are different from the situations on land where the communication takes place with radio frequency (RF) aid. Unfortunately, the underwater environment consumes the energy of the RF waves and renders itself impractical. The mobility of underwater sensors with the ocean currents is another major challenge. To get better off from the challenging underwater situations, UASNs communicate using acoustic waves [8–10]. Underwater acoustic waves typically operate in the frequency range of 10 Hz to 1 MHz. The delay of propagation accompanies this slender range, but it seems to be the only viable choice forward on modicum energy store. UASNs make another edge by facilitating interfaces to communicate with autonomous underwater vehicles (AUVs) remotely. This feature will exponentially increase the range of AUV control, and this merit will also help us perceive the underwater world for research with an added advantage of burgeoning the amount of ocean monitored by human beings, which currently accounts for only 5%. Routing of data packets right from the sensor nodes to the sonobuoys and to surface stations is one of the most challenging issues faced in UASNs, primarily due to the rapid energy drainage, limited bandwidth, significantly high latency, and reduced reliability [11–17]. Figure 1 presents a sample application scenario of UASNs.

Terrestrial wireless sensor networks (TWSNs) have a conventional set of routing protocols that ensure good network performance. TWSNs at no point of operation face interruptions similar to ocean currents. Doppler spreading, interim path loss, and link quality loss create numerous challenges for routing in underwater environments. In a nutshell, the quality of service (QoS) and energy constraints of UASNs inextricably impede it from resorting to routing protocols of TWSNs. Majority of all the routing protocols proposed for TWSNs, thus proving to be powerless when it comes to UASNs. Numerous unconventional routing protocols were put forth in recent years for UASNs, and each focused-on energy efficiency, thus improving various QoS parameters like throughput, latency, load balancing, and robustness. Some of these protocols have already been tested for research, military applications, and catastrophe prediction. The selection of an appropriate routing protocol is significant as it is answerable for the reliable deliverance of data packets to the destination.

Table 1 presents the variations between the terrestrial wireless sensor networks (TWSNs) and underwater acoustic sensor networks (UASNs). Routing protocols in UASN face numerous design challenges. The weightage given to path selection accounts for the various problems that have to be confronted in the underwater environment, such as marine aquatic life, acoustic disturbances, propagation delay, and seismic shadow zones. Many new routing protocols are proposed to tackle these dilemmas; however, most of them lack the description of appropriate routing strategies. Routing strategies advocate the parameters, which will be extensively useful for researchers and other professionals to calibrate the effectiveness of algorithms used in UASNs to develop a strategy to tackle limitations like high propagation delay and energy usage. Picking the suitable scheme ensures engineers achieve desired productivity in applications. The routing protocols for UAWNs are mainly classified into location-based protocols and location-free protocols. The location-based protocols instrument the use of the information contained in the sensor nodes that are mostly two/three-dimensional position coordinates. In contrast, location-free/depth-based protocols depend mainly on pressure information present in sensor nodes. Most of the earlier conventional protocols proposed for UASNs selects the best path for sending data beforehand without considering the dynamic nature of the network environment. This negligence of the traditional routing protocols tends to compromise the use of the widespread resources in the network and can also pave the way to network failure. When setbacks like these started to portray, routing in UAWNs seems like an insurmountable dilemma, and the concept of opportunistic routing protocol (ORP) was then proposed [18–20]. This traction appeared due to various contributing factors like the increased need for extended capacity and expectation of top-notch QoS. The basic idea behind ORP turns the table on the demerit of unreliable transmission, that is, the undesired broadcast nature exhibited by the unreliable transmission is exploited here instead of selecting the nodes beforehand. The selection of nodes in ORP happens on the go. Numerous neighboring nodes (candidate set) receive the broadcasted message. The candidates belonging to the candidate set are sorted according to the metrics and prioritized based on the probability of becoming the next-hop forwarder. The candidate with the highest priority is given the ability to forward the data packets while others discard the packets. This is known as candidate coordination. Opportunistic routing protocols have proved their robustness and adaptability to uncertain conditions by showing their significant presence in many essential fields like oil/gas pipelines, power grids, and management of metro/railroads. Currently, the advancements closely related to ORP have not yet reached their pinnacle as many problems are yet to be solved. However, the prime intention that sleeps behind it is being the ability to make a set of independently weak nodes emerge together as a virtually robust set of links. Thus, ensuring reliability which in turn plummets the retransmission rates and chop down the energy consumption of UAWNs. All of these pros and cons will be thoroughly surveyed in this study.
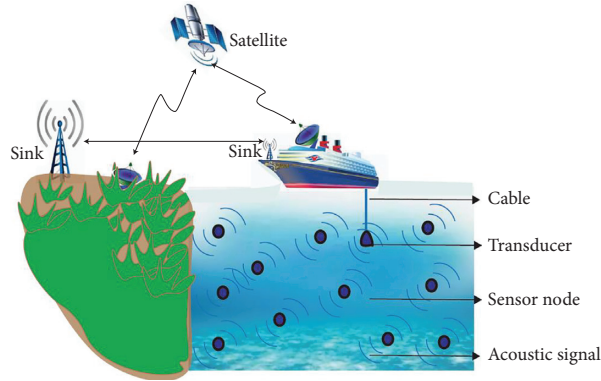
FIGURE 1: Underwater acoustic sensor networks (UASNs).

TABLE 1: Differences between the TWSN and UASN.

	TWSN	UASN
Mobility	Low/medium	High
Reliability	High	Low
Data rate	High	Low
Bandwidth	High	Low
Interference	Low	High
Propagation delay	Low	High
Energy usage	Low	High

The major contributions of the work are highlighted as follows:

(i) We review all the major location-based opportunistic protocols proposed for routing data packets in underwater acoustic sensor networks over these years. Numerous energy-based, QoS-based, and hybrid location-based opportunistic routing protocols have been proposed in 2019, 2020, and 2021, and they promise to provide much better performance to various real-time applications deployed in UASNs. Very few works have provided reviews on these latest protocols, and we initially tried to address this research gap. We analyze and present a brief description of their working mechanism and highlight their issues and challenges. These issues can be taken up in the future for further improvement in the design of routing protocols in UASN.

(ii) From the review of the latest protocols, it is observed that increased delay and energy drainage are the two significant areas of concern that need further solutions. We then tried to model an easy-to-implement routing protocol that can guarantee energy efficiency with reduced delay to various applications deployed in UASNs.

The rest of the study is organized as follows: in Section 2, various applications using underwater acoustic sensor networks are discussed. The fundamental principles of opportunistic routing are explained in Section 3. Energy-based, QoS-based, and hybrid location-based protocols are discussed in Section 4. The proposed energy-efficient and delay optimized protocol is discussed in detail in Section 5. The future research directions are discussed in Section 6, and the study concludes in the next section.

2. Underwater Acoustic Sensor Networks

The planet we dwell upon is covered approximately with 71% of water. Under this blue element, lies countless untapped resources that will enable human society to advance in countless ways. In order to consolidate this final frontier, underwater wireless sensor networks prove to be the need of the hour. The underwater wireless sensor network (UASN) is the collection of self-driven sensor nodes and autonomous vehicles connected underwater to perform different collective tasks based on user applications. The sensor nodes can occupy different depth locations that will permit us to spread our reach even to the ocean's deepest places. The self-driven sensor nodes will collect and transfer the sensed data to the target destination using acoustic signals. The attractive applications of UASNs comprise real-time surveillance, disaster prevention, navigation assistance, water quality determination, industrial organization, fish tillage, underwater exploration, and pollution tracking. Almost each of these applications of UASN demands sensor nodes to transfer sensed data timely and precisely through the source node present underwater towards the destination node on

the surface with the help of intermediate nodes in the network. However, due to the dynamic nature of UASNs, continuous node mobility, communication voids, and limited battery storage often lead to degraded network performance. In this complicated underwater environment, how to route data packets promptly and effectively even with the presence of a communication void is the most challenging research question.

2.1. Challenges in UASNs. This section presents an overview of various challenges in underwater acoustic sensor networks.

2.1.1. Acoustic Communication. The terrestrial networks depend on RF waves to execute communication. Whereas the underwater environment is highly unpredictable, and RF waves are unfortunately absorbed. Additionally, a high amount of attenuation paves the way to energy loss; hence, RF waves are not an option to be considered. Optical waves cannot be regarded as a choice because the mobility of the nodes is unable to guarantee accuracy. The only viable option in this scenario is acoustic waves. Many other flaws are associated with the usage, but a suitable routing protocol is expected to sort out the dilemma.

2.1.2. High Mobility. The underwater sensor nodes are constantly on the move. These movements are caused by ocean currents which arise due to wind, breaking waves, temperature, and salinity variations. For efficient data gathering, the movement of these nodes is vital and indispensable. However, in reality, this high mobility induces the formation of curves to the acoustic waves, which triggers the emergence of zones that makes some of the sensor nodes in the network unable to participate in data transfer, which brings forth performance issues to the network.

2.1.3. Underwater Noise and Interference. The ocean is packed with a lot of noises and interferences that arise from varied sources. Some of the underwater noises are caused by breaking waves, rain, and marine life. However, various sources are man-made like, shipping, military sonars, fishing, and research activities. These disturbances affect the quality of data packet delivery in the underwater environment.

2.1.4. Low Bandwidth. The operational frequency range of the underwater sensor nodes is primarily restricted due to the usage of acoustic waves as the medium of transmission. The bandwidth is a meager spectrum that lies between 1 kHz and 50 kHz. This poses a significant problem for routing protocols as it requires an enormous amount of data exchange at various stages like discovery and maintenance. Tight bandwidth constraints put challenging design constraints on acoustic systems. In order to perform communication with AUVs, it is more important to have a wider bandwidth rather than a rate. Moreover, the routing

protocols are forced to choose routing paths from this small frequency range for data delivery. However, due to the dynamic nature of UASNs, continuous node mobility, communication voids, and limited battery storage often lead to degraded network performance. In this complicated underwater environment, how to route data packets promptly and effectively even with the presence of a communication void is the most challenging research question.

2.1.5. Low Data Rate. Speed is a crucial factor when it comes to information exchange. The faster the data reaches the destination, the better. Unlike its counterpart (terrestrial environment), the speed at which data is transmitted in the underwater environment is influenced by numerous factors. Firstly, the propagation speed of acoustic waves is inferior to RF waves by many folds that create room for propagation delay. Secondly, there are various persuasive components like depth, temperature, and the degree of salinity of the water. The data rate is meager and accounts for approximately 100 kbps or occasionally a bit more.

2.1.6. Transmission Loss. The hurdles when it comes to underwater sensor network implementation are numerous. Acoustic waves do not guarantee any reliability for the network. On top of that, path loss, Doppler spreading, and high latency will provide a clear picture that there is a considerable amount of packet loss. Transmission loss in any network is not a desirable outcome. Interference is one of the main reasons for packet loss to occur. When the same nodes participate in data transmission continuously for an extended period, the battery can run out, resulting in a communication blackout, and the node will no longer be able to transfer data collected from some places in the network. The acoustic signals have open channels, which are more likely to be utilized by an attacker or malware and to wreak havoc in critical services like routing, localization, and synchronization. Delay variance and bit error are two constituents that can lead to a high amount of packet loss and bit error rates.

2.1.7. Error Prone. The underwater sensor nodes, unlike their conventional counterparts, are not reliable. The environment in which it has been implemented does not support its operation. The mobility, high latency, delay in propagation, high interference, noise, etc., make them highly susceptible to errors. The changes that manifest due to the variation in salinity, depth, and acoustic speed have an implicit effect on making the data transmission process error prone.

2.1.8. High Energy Consumption. Energy drainage is a significant problem in UASNs. Communication between various nodes in a network rudimentarily requires the acknowledgment of its position. The nodes are constantly swaying in harmony with the ocean currents, and it is essential to update their position consistently with their neighbors for effective participation in the data transmission

process. Ironically, this position-update process drains quite some energy from these sensor nodes. Another avenue wherein the power consumption rates rocket is when packets have to be retransmitted due to high interference. The data load bestowed on end nodes that connects to the surface stations rapidly depletes the battery power, resulting in the termination of connection with the rest of the nodes in the underwater network. The consumption rates vary with the depth in which it is implemented; the battery is expected to operate well in shallow and deep-water conditions.

2.1.9. Channel Attenuation. Channel attenuation is another dilemma that underwater sensor nodes have to confront. The implementation of sensor nodes in the ocean bed is beneficial only if it can collect and transfer data, but due to channel attenuation, the data collected cannot be efficiently extracted from the received signal.

2.1.10. Short Network Lifetime. The pivot grounds for the short lifetime of underwater sensor nodes are due to its source of energy, the battery storage. The nodes extensively consume energy while localization, routing, and data transfer. The hostile underwater condition makes the replacement of cells regularly a challenging task. Therefore, an efficient routing protocol has to consider energy consumption while making decisions on routing path selection. Furthermore, if the nodes run out of energy, it will result in the formation of dead nodes that can affect the network performance and data transmission to the surface stations.

2.1.11. Security and Privacy Issues. UAWNs are made and deployed to monitor places consistently that are far from the shoreline. There is a high chance that the nodes are deployed in strategic waters for specific applications. Attackers can easily manipulate UASNs to inject malicious attacks upon the network. They can also be physically destroyed by enemy divers/AUVs. In the worst-case scenario, attackers can inject fake nodes into the network to provide misguided information and use compromised nodes to extract exclusive data from the network.

2.1.12. Physical Challenges. The sensor nodes have to be fabricated so that they are compact, rigid, and waterproof and should also be able to withstand the pressure of water. Marine life is the next physical challenge that these sensor nodes have to face. In reality, it is physically impossible to protect every one of the nodes in the network.

2.2. Applications of UASNs. This section presents an overview of the major applications that use UASNs.

2.2.1. Military Applications. The military applications of UWSNs can cover a wide range of requirements from monitoring to reconnaissance. In 1982, the United Nations Convention On the Law of the Sea allowed countries to exercise jurisdiction on territorial waters up to 200 nautical miles along the baseline. The naval force guards the coastline against invaders, but the underwater regions in strategically important areas are vulnerable. This vulnerability can be defeated with the help of UASNs. It will enable the militaries to detect enemy divers, submarines, torpedoes, AUVs, and naval mines. The ability to get real-time data will enhance strategic decision-making.

2.2.2. Oceanography. Oceanography is the study of physical and biological aspects of the ocean. Oceanographic studies and researches will provide humans with the capability to understand the various phenomena that take place in the ocean and also be able to predict or artificially simulate similar conditions that will benefit human society. Oceanography can provide efficient analysis if it can collect real-time data. UASNs can be utilized to perform experiments to unravel the mysteries of the underwater world consistently, which will indirectly help us to obtain solutions to various problems oceanographic problems.

2.2.3. Coral Conservation. Corals are one of the most beautiful living things on the planet. It takes millions of years to form barrier reefs. The Great Barrier Reef in Australia is the most extensive collection of corals on the phase of the Earth. However, it is dying due to coral bleaching due to the shift in climatic conditions throughout the world. Coral reefs around the globe are on the verge of extinction, and nations are trying to change the situation. Human interference has the likeliness to accelerate the degradation process, but data collection is seemingly impossible without human interference. UASNs are a profound solution that can provide real-time data to conservationists by limiting human interference.

2.2.4. Resource Tapping. The Earth has resources that are tucked away in the depth of the ocean. UASNs provide a way to get the know-how of these varied resources. UASNs will enable us to find out the location, approximate quantity, and dispersion pattern of resources present in the ocean bed. Petroleum and natural gases are an inevitable part of our civilization, which are unfortunately limited. New potential sources can be located for extraction using UASNs.

2.2.5. Fishing, Farming, and Recreation. The fishing industry will benefit from the use of UASNs as they will help locate groups of fish. Underwater farming has been used to cultivate seaweed, lettuce, basil, etc., in countries like Japan and Italy. Knowing the nature of the ocean is an integral part of taking out cruises for recreation purposes deep into the ocean. UASNs can tell leisure seekers about the risks of a tsunami or hurricane before they embark on a cruise.

2.2.6. Disaster Prevention/Prediction. The implementation of UASNs will enable us to detect in advance any underwater earthquakes and volcanic eruptions, which will help us prevent or predict disasters. Many aircraft have gone missing

over these years in oceans, and no data about their disappearance was harnessed. One of such shocking incidents is that of Malaysia Airlines 370. UASNs, provided it is implemented correctly, will enhance us to chart the ocean, and it is possible to derive a pinpoint location of any possible crash site.

2.2.7. Climate Change. The rising sea levels and warmth have grabbed international attention. The polar ice caps are at the risk of meltdown. Accurate screening and reports of the polar meltdown can be undertaken with the help of UASNs, which will help researchers and environmentalists to find solutions to these baffling dilemmas.

3. Opportunistic Routing in UASNs

The uncertainty of the underwater environment poses many underlying threats to establishing efficient communication strategies. The constraints on power and the constant movement of nodes due to tides make opportunistic routing (OR) a viable solution. The opportunistic routing owns different modus operandi. On receiving a data packet, the host node takes into consideration a set of eligible neighboring nodes and prioritizes them based on various parameters. These parameters are different facets like the node closest to the destination and least power draining. A suitable packet forwarder gets opted from the candidate set based on priority and availability. Thus, in the case of unreliable underwater communication, OR proves to be promising as it provides extended reliability, robustness, and QoS than other legacy routing methods.

The principle of opportunistic routing idea was initially developed in ExOR [19] in 2005. The notable advantage of this protocol is that this protocol exploits the multiplex communication opportunities in which the broadcast character belonging to the wireless network develops. The fundamental working of opportunistic routing can be implemented in this protocol, and the three main steps include the following: Initially, the sender node can broadcast the message data. Secondly, upon receiving that data, one relay node is selected as the best forwarder node. After that, the selected best relay node transports the message transmission to the next best relay, and so on. The method is continued until the data reaches the target position. Compared to traditional routing methods, the next-hop relay is selected only after it has received the data, thereby reducing the number of data retransmissions.

The UASN's operation using acoustic channels for communication has many downfalls as there is a prevalence of solid attenuation, time-varying multipath, ambient noise, and modicum propagation speed. All of these contribute to increased delay, error, limited bandwidth, high energy consumption, communication cost, and at times temporary loss of connectivity among nodes of the network. The profound influence of channel fading is crucial to confront as it can directly impact declining routing performance. The scope of application of OR is of paramount importance because it has significantly low retransmission rates, in turn

reducing the power consumption. Assured packet delivery facilitation by opportunistic routing ensures no wastage of network resources. It also reduces the chance of system collapse and diminishes retransmission costs. Additionally, opportunistic routing is a versatile choice as it applies to a variety of networks.

Opportunistic routing facilitates a dynamic and instant multiple-path routing technique through opportunistic relay selection, unlike the traditional routing method. Instead of a single precomputed relay, opportunistic routing initially broadcasts a data message to a set of forwarder relay nodes. Fundamentally, these forwarders are organized according to a particular unit. The idea of this method is to select the best forwarder relay among all the nodes. This selected relay manages the process of forwarding the packets. These steps are recursively executed until the packet is transmitted to the target destination. The rudimentary opportunistic routing mechanism consists of the following four steps:

(i) Forwarder relay set choosing

(ii) Broadcasting of data to forwarder nodes

(iii) Coordination scheme is used for best relay selection

(iv) Forwarding of data by the best relay

Each node in opportunistic routing broadcasts a data packet to various adjacent hops periodically. Hence, if communication to one neighbor crashes, another nearby node that has received the data packet can transmit it. OR defines a group of various next hops as the best forwarder relay collection and indicates it as a forwarder relay set (FRS). When a message gets transmitted to FRS, numerous forwarders can obtain a similar message packet. We can avoid duplicate transmission by selecting a single candidate as the best relay. Each node in the FRS is allocated a priority that is calculated based on a predefined variable. Suppose the node with the most significant priority present in the FRS favorably acquires the message, it transmits that to the target. Otherwise, the node that has the next most significant priority transmits the data packet, and so forth. The leftover candidates will discard this data packet. The FRS selection is divided into three main components as follows: (a) forwarder relay discovery, (b) prioritization variable calculation, and (c) forwarder relays selection, prioritization, and filtering.

To find out the neighbor node, periodic or nonperiodic packets are broadcasted. This neighbor node depends on the link quality, which is changeable as well as dynamic. Hence, this phase is given the charge of computing the stability and quality of the links to reach the neighborhood. Based on these values, a group of nodes is determined. Initially, every node located in the vicinity of the sender node is added to the FRS. Then, the forwarders are taken and sorted based on the chosen variable. After an FRS is elected, priorities are given to forwarder relays based on a specific value. So, the variable selection affects the network throughput significantly. The priority variable election depends on the routing application needs and targets. For some applications, such as emergency recovery, the position data is essential. Consequently, the relay nodes must know their location, and routing can be

executed by selecting the relay that is locationally nearest to the target node. Controlling the number of forwarder relays can reduce the overhead and duplicate data transmissions. Moreover, since the size of FRS grows, the number of forwarders who cannot listen to one another also increases. This leads to duplicate packet transmissions. Hence, it is suitable to avoid some forwarder candidates from the FRS. This approach is termed candidate filtering or forwarder. The conventionally used filtering method avoids the forwarders, which are not suited instead of the source. But somehow, this policy cannot guarantee efficient network performance. Another method of FRS creation is discussed in some previous works. The technique is based on implementing an algorithm that can optimize correctly and compute the optimal forwarder relays set for every node, such as forwarder relay sets that are created by Dijkstra's algorithm. However, these methods do not solve the issues of duplicate data transmission.

The optimal relay selection that uses a coordination scheme is used for the coordination of packet forwarding operation between next near nodes. This scheme is responsible for selecting the most suitable forwarder relay to push forward the data packet. Coordination methods need signaling between forwarders. The basic coordination schemes are generally classified into a timer-based, contention-based method, and token-based coordination scheme. In a contention-based method, the main principle is that forwarder relays contend to transmit the data packet with the help of control messages. For example, if a sender node transmits a forward request, its near hop nodes have to compete with themselves to come to a consensus on the forwarding of the data packets. In the timer coordination method, the forwarder relays are supposed to be ranked according to a specific priority value. This rank is commonly added with the message header, which is consistent with the hierarchy in which potential forwarders are permitted to respond. Thus, the largest priority node is allowed to respond to the first slot. The next priority node responds to the upcoming time slot, etc. However, this method is straightforward and easy to carry out, timer-based coordination that incurs some delay, affecting network performance. Another method is the token-based coordination scheme, in which the transferring of data packets is only possible through a token holder. In this scheme, the duplicate message transmission is completely prevented, but it faces increased overhead control. A forwarder (relay) node contains the overhead data packets that are being sent when a token arrives. Tokens travel with connected forwarders because more miniature priority forwarder relays can listen to large priority nodes. If no token arrives, the candidates may be moved into an idle state, slowing down the network.

4. Location-Based Opportunistic Routing Protocols in UWSNs

All routing protocols for UAWNs can be classified as location-based protocols and location-free protocols. The location-based protocols instruments data contained within sensor nodes that are mainly two/three-dimensional position coordinates. At the same time, the location-free protocols/depth-based protocols depend on the information related to pressure on sensor nodes. The focus of this study lies in location-based protocols as they provide better performance than location-free protocols. Vector-based forwarding (VBF) [21] was one of the earliest protocols proposed in this category. This protocol can fabricate some virtual vector pipe that exists between the source and the destination. Only the nodes in the vicinity of the "vector" right through the source and destination will have the ability to do message forwarding. Hence, routing involves only a tiny group of nodes.

Similarly, numerous location-based protocols were proposed for UASNs. Some of the protocols focused on improving energy efficiency, while others focused on improving QoS parameters like delay and delivery ratio. Recently, many hybrid protocols also have been proposed which consider both energy efficiency and QoS. This section presents a comprehensive discussion on all the latest location-based protocols proposed for UASNs.

VBF makes use of the node location information to make the routing decisions. The knowledge of the position of nodes encourages it to be faster, reliable, and scalable. With the protocol, a virtualized pipe is created from source to destination, and those nodes in the pipe possess a higher probability of becoming the forwarder nodes, while the nodes outside are disregarded. Sink-initiated query and source-initiated query are the two main ways in which VBF addresses routes to different queries. Conceptually, all nodes inside the virtual pipe have the eligibility to forward the packets, but due to limitations like energy, mobility, and propagation delay of acoustic waves, a self-adaptation algorithm was suggested. Another protocol, directional flooding-based routing (DFR) [22] defines a forwarding method formulated by the angle among the center and intermediate nodes. The nodes are responsible for forwarding the data packets through a flooding method. It also considers the quality of the link between the sender and the destination node. A significant concern with this protocol is redundant data transmission and increased energy consumption. The information-carrying routing protocol (ICRP) [23] is an influential conservative, continuous, and versatile directing protocol. The sender hub checks the current location to the final destination when it owns the data to be sent. If there is no current course or path, it starts a path development process by communicating with the information packet, conveying the route disclosure message. Every node present on the network communicates to maintain the reverse route with the information path.

Hop-by-hop vector-based forwarding (HH-VBF) [24] is another variation of VBF where each forwarder resorts to a different routing vector. HH-VBF rudimentarily is just a version of the vector-based forwarding protocol. HH-VBF is a viable option compared to VBF as it can work well with sparse networks and is not liable to the routing pipe radius threshold. However, there is an increase in the computational delay, which in turn degrades the network performance. Reliable and energy balanced routing (REBAR) protocol [25] is a routing protocol that is energy efficient that

helps in varying the broadcast domain. REBAR has good reliability and increased lifespan of the network. To balance the energy consumption in the network, a flexible scheme is developed to establish the data propagation range. Here, the nodes near the destination have a modicum radius. However, the increased node movements may expend an excessive amount of energy, resulting in degradation of performance. Vector-based void avoidance (VBVA) [26] protocol is simply an extension of the VBF protocol and focuses on addressing the void problem and energy efficiency in UASNs. The protocol works similarly to VBF when there is no void but uses a revised strategy when voids appear in the network. This helps the protocol to maintain better energy efficiency even with voids in the network.

The energy-efficient and collision aware (EECA) [27] multiple-path routing methods are founded on computing two different collision-free paths using restricted-energy modified flooding. It is one of the earliest protocols that gave equal importance to the betterment of QoS and energy efficiency in the network. Here, multipath power-control transmission (MPT) allows packet data transfer within limited end-to-end data error value and reduced power of transmission. The reliable energy-efficient routing protocol [28] functions on the foundation of link quality, physical data distance, and energy available in the UASN. These three values are calculated and shared with all nodes in the network. The protocol uses a local flooding mechanism with an adaptive selection and gives good reliability and energy efficiency performance.

Location-aware routing protocol (LARP) [29], the GPS is used to identify the exact area of the sink nodes. The sink nodes then broadcast the location information in the network. At the least three sink nodes are used for reference, other nodes in the network calculate their position. The sender can locate the next hop by broadcasting two things as follows: (1) location of the destination node and (2) moving direction of the packet. Packets are forwarded if the receiving node discovers that it is moving in a similar direction. The quality-of-service aware directional flooding-based routing (QoSDFR) [30] extends the DFR protocol. In this routing strategy, the sink node is responsible for sending feedback to various other nodes in the network about the channel condition, and based on the feedback, the optimal forwarder is selected. Protocol results in high throughput because of the limited energy consumption and varying channel conditions.

Scalable and efficient data gathering (SEDG) [31] protocol tries to increase the delivery ratio of the packet and also saves the modicum energy by feasible assignment of the member nodes and gateway node (GN). Here, an autonomous underwater vehicle (AUV) goes through the network area with a precomputed elliptical route and collects data from the gateway node (GN). AUV-aided efficient data gathering (AEDG) routing protocol [32] employs an AUV to gather information from gateways or intermediate nodes and use the shortest path tree (SPT) algorithm to balance the energy consumption. Besides that, AEGD designs a model that improves the result and saves energy by reducing the node members. Furthermore, the nodes live for an extended

time to transfer data, thereby increasing delivery chances. The delay-aware energy-efficient routing protocol (DEEP) [33] is a delay-aware routing protocol dependent on collision rate and energy. DEEP makes use of an adaptable node aimed to minimize the collision rate. All the intermediate nodes are elected by virtue of delivery ratio and link quality. In the channel aware routing protocol (CARP) [34], the next-hop transmitter node is elected due to its distance from the previous intermediate node and available energy. Here, every intermediate node is familiar with its neighborhood between the destination node and the next hop. Sender then broadcasts a PING message to the network to compute the next forwarder. Considering a case where the hop value of a sink is lower than the sender node, it replies a PONG data. CARP uses an efficient relay selection method, which doubles the packet delivery ratio.

The novel efficiency forwarding protocol (NEFP) [35] is a proactive anycast routing protocol proposed for UWSNs. It promotes three different approaches. One defines a routing method that avoids unnecessary forwarding of packets where the collision dilemma is averted using a timer. Moreover, finally, the design uses Markov chains to calculate the probability of forwarding the data packets that encourages adaptability to constantly changing network topology. Nevertheless, the performance of the suggested protocol is decreased in the sparse region and as a result, reduces the number of forwards in the phases. Geographic and opportunistic routing protocol with depth adjustment (GEDAR) [36] is a geo-opportunistic routing protocol proposed for a minute-monitoring task. It utilizes a greedy forwarding method to advance the message towards the next hop. The source node chooses the best candidate from the forwarding set. The opportunistic routing in GEDAR reduces the number of retransmissions. GEDAR uses a recovery mode that helps to avoid the void areas. If a node is present in the void area, it will adjust its depth to overcome the void, and new messages will be queued. The greedy strategy will reschedule the node later. Markov model-based routing (MMVR) [37] selects its route from the lower surface to the top level based on changing data traffic. The routes are stable and adaptable, with fewer hops from the sender node and destination. In a localization-based dynamic routing protocol (LBDR) [38], the network is split into smaller layers, and a virtual routing vector is made within the sub-layers. Nodes will move in and out of the virtual vector based on the water current, resulting in high throughput. Void handling geo-opportunistic routing (VHGOR) [39] protocol focuses more on efficiently handling the communication holes in the network. Here, a quick hull algorithm is used to avoid a convex hull. When the node or hub reaches a convex region, rebuilding the convex void helps check an alternate and different way to resume the greedy transmission. VHGOR improves the network performance in networks with voids compared to other routing protocols.

Geographic and opportunistic routing (GOR) [40] protocol shows efficient multi-hop data transmission in UWSNs with an upgraded strategy compared to previous protocols. Sometimes, this method gives room for the

formation of the void region, and GOR tackles this issue using some void-handling algorithms. The framework considers network density, traffic load, and energy control features to bypass the empty region. The range-based low overhead localization technique (LOTUS) [41] significantly improves on earlier versions of the localization protocols. The protocol can estimate locations based on only two references, enabling this technique to work in networks with fewer nodes. The geographical duplicate reduction flooding (GDflood) [42] considers the location data regarding sensor nodes and joins it with network coding. Energy-efficient grid routing based on 3D cubes (EGRCs) [43] employs a 3D cube network that is subdivided into small cubic clusters. The cluster head is determined based on the remaining energy and position of the intermediate node. All the cluster heads then compute their intermediate node based upon the delay and location. EGRCs reduce energy consumption and end-to-end delay and increase the network performance.

Mobile energy-efficient square routing (MEES) [44] is a routing protocol focusing on energy efficiency in underwater sensor networks. The method uses a division of the network field into dense and sparse regions. A major advantage of this method is that, the mobile sink shifts in a clockwise direction that ensures the highest coverage of nodes in the network which will, in turn, result in high throughput and energy consumption. Topology control vector-based forwarding (TC-VBF) [45] is a revamped version of VBF, which tries to address the limitation of VBF in light conditions. Another protocol energy-efficient multipath grid-based geographic routing protocol (EMGGR) [46] fragments the network into 3D grids. The routing is executed in a grid-by-grid fashion with the help of gateway nodes. Disjoint paths result in high energy efficiency and a good packet delivery ratio. Balanced multiobjective optimized opportunistic routing (BMOOR) protocol [47] uses a strategy where the data from the lower surface takes the best route through the intermediate nodes to the top-level sink. Here, the nodes are located as per dynamic assessment with regards to optimal energy forwarders. The BMOOR protocol needs no spatial data, which is costly in UWSN. The protocol is developed using a generation-based bio-inspired, meta-heuristic algorithm. This helps in delay depreciation and maximization of delivery ratio, and thereby the network lifetime is enhanced. Another proposed protocol for UWSN is energy-efficient interference aware routing (EEIAR) [48] that opts for the best forwarder following the shortest distance. The shortest distance determination decreases the propagation delay. The power control-based sharp directing routing (PCR) [49] selects the most optimal transmission power level available at each submerged sensor node, which helps improve the packet delivery conveyance at each round. Also, it condemns the usage of high-power transmission and the uncontrolled consideration of neighboring hubs in the following hop candidate set, which would end up being the root cause for building the energy utilization on the network. The simulation outcomes depict that PCR diminishes the energy expenditure by adjusting the transmission power and electing the best candidates. The stateless opportunistic routing protocol (SORP) [50] uses a novel method to employ a variable forwarding area that can be reshaped and replaced according to the regional density and placement of the potential forwarding nodes to improve the energy and reliability. The protocol gives good performance compared to the previous protocols. Glider-assisted link disruption restoration mechanism (GALDRM) [51] uses a link disordering recognition with a related link rebuilding method. In the connection acknowledgment system, the group nodes gather the link data. The cluster heads gather the disruption data in link disruption and then schedules gliders as relay nodes to revive the link. Utility capacity is built up by limiting the channel. A multiplier technique illuminates the ideal area of a lightweight flyer. The simulation outputs exhibit a glider-assisted reconditioning procedure that helps to reduce energy consumption. The energy-aware void-avoidable routing protocol (EAVARP) [52] expands the network lifetime and packet delivery rate in underwater sensor networks. EAVARP includes layering and data collection phase with the help of directional forwarding strategy and uses residual energy and data transmission to avoid cyclic transmission and flooding. Fuzzy logic-based VBF protocol (FVBF) [53] improves VBF protocol. It focuses more on the selection of a single forwarder node in VBF. FVBF is the fuzzy logic-based VBF protocol. The best forwarding node is chosen according to the angle of projection distance and the battery level. The smallest distance shows that the node is in the vicinity of the target node. The projection angle allows it to be selected onto the virtual routing vector pipe. The best advantage of this protocol is that it achieves better energy and throughput. However, nodes in the selected vector terminate on dealing with a high load of the message, which is similar to the conditions in VBF.

Mobility-assisted geo-opportunistic routing (MSAGOR) [54] protocol is mainly based on interference avoidance. Here, the network region is fragmented into compact cubes to diminish the interference, which helps to make additional well-informed routing strategies for better energy utilization. Moreover, an optimal number of transmitting nodes are selected from each cube based on its distance to the destination. This proximity will help to avoid void nodes. The extensive simulation results reveal that this protocol will maximize the delivery ratio and network lifetime. Totally, opportunistic routing algorithm (TORA) [55] is an anycast, geographic opportunistic routing protocol proposed for UWSN. The protocol is implemented to avert parallel transmission, bring down end-to-end delay in the network, tame the dilemma of void regions, and enhance network throughput. TORA uses time on arrival and its range-based equation to localize nodes. The energy-aware opportunistic routing (EnOR) [56] is an energy-aware opportunistic routing (EnOR) protocol that can adjust the priority level of forwarding between candidate nodes. This leads to steady energy utilization and increased network lifetime. By using the residual energy, link reliability, packet advancement ratio, and EnOR change the priority of transmission level. Adaptive hop-by-hop cone vector-based forwarding protocol [57] tries to improve the reliability of data transmissions in the sparse sensor regions by making some modifications to the base angle of the cone as per the

network structure. These protocols improve the network performance by reducing the number of duplicate packets and also enable a better selection of the potential forwarder node.

Authors in reference [58] discuss implementing a modified strategy for depth-based routing that can transfer the data reliably to the surface sonobuoy. The technique mainly uses the 2-hop neighbor technique and tries to improve the delivery ratio of packets in the network. Authors in reference [59] proposed a technique combining the ant colony optimization algorithm, artificial fish swarm algorithm, and dynamic coded cooperation to improve efficiency by reducing energy consumption. Improving the flexibility of the protocol with the network was one of the major tasks of the proposed algorithm, along with finding the most optimal route. In reference [60], authors presented a Q-learning-based multi-hop cooperative routing protocol for underwater networks. Using this algorithm, the nodes with maximum Q-value were selected as the next forwarders in the network to transfer data from the source to the destination. A coding-aware strategy was proposed for efficient routing in networks with the sparse deployment of nodes [61]. The topological information was used to expand the candidate set using the protocol. An interesting approach that utilizes AUVs to carry sensor nodes to repair the routing voids when foreseeing the occurrence of voids was proposed in reference [62]. The protocol initially predicted the location for repair and then directed the AUVs to the particular location to carry out the repair process. Most of the proposed protocols are complex and incur high overhead, which degrades the performance of the network. Although many of the current protocols improve the data delivery ratio significantly, it comes at the cost of increased energy consumption. It is vital to develop a simple to implement a protocol that can take care of energy efficiency in the network while ensuring reduced delay in the network.

5. Energy-Efficient Location-Based Opportunistic Routing Protocol (EELORP)

In this section, we present the discussion on the proposed energy-efficient location-based opportunistic routing protocol (EELORP) that is designed to provide better energy efficiency and data delivery with minimum delay.

5.1. Theoretical Analysis. Initially, we try to provide a theoretical analysis to the proposed protocol. The focus is mainly on the delay of transmissions that can be reduced further to enhance the performance of the system. In the underwater network, a delay occurs within two different links, the wireless sensor to the wireless controller and wireless controller to the actuators. The delays are denoted by $T_{w(s-c)}$ and $T_{w(c-a)}$. Assuming the controller to be time invariant, the delays due to two sources are combined together to get total wireless sensor network delay as follows:

$$\text{Twt} = \text{Tw}(s-c) + \text{Tw}(c-a). \tag{1}$$

The computation delay of the controller can also be included in the total wireless sensor network delay. As the assumption in the wireless controller is time invariant, the decision of controller $d_{(t)}$ is independent of the time it receives the sample $S((\gamma b)$. So, the total wireless sensor network delay is only important for us. The analysis of UWSN stability is carried out by assuming two different scenarios as follows: (a) the continuous UWSN network delay system is considered by determining UWSN stability with constant network delay and (b) the discrete UWSN network delay system is considered by determining UWSN stability with time-varying networking delay.

For a continuous UWSN network-delayed system, the UWSN having total network delay as T_{WT} at the time $t = \gamma b$ is considered. The assumption is extended by making $T_{wt} < b$ for all values of "γ" belonging to "S" with $\gamma \in s$. The system is modeled mathematically as follows:

$$\frac{\partial}{\partial t}l(t) = Pl(t) + Qd(t - T_{WT}), \tag{2}$$

where "t" belongs to $[\gamma b, \gamma b + b]$. Also with P, Q, R, and S as known matrices, we have the following relation:

$$M(t) = R\gamma(t) + Sd(t). \tag{3}$$

$d(t)$ is the received signal with no delay and $d(t - T_{WT})$ is the received signal with delay. In the case of $d(t)$, that is, received signal with no delay $d(t) = d(\gamma b)$ for $t \in [\gamma b, (\gamma b + b)]$.

Proposition 1. *The UWSN with the delay mentioned above validates the below-mentioned difference equations. The derived equation is as follows:*

$$l(\gamma b + b) = \tau l(\gamma b) + \varepsilon_0(T_{WT})d(\gamma b) + \varepsilon_1(T_{WT})(d(\gamma b - b)). \tag{4}$$

For $\tau = \int_0^b e^{Pb}$, we obtain the following equation:

$$\varepsilon_0(T_{WT}) = \int_0^{b-T_{WT}} e^{Pb}d\gamma\varphi,$$

$$\varepsilon_1(T_{WT}) = \int_{b-T_{WT}}^{b} e^{Pb}d\gamma\varphi. \tag{5}$$

Now, we have the following equation:

$$l(t) = e^{P(t)} + \int_0^t P(t-r)\varphi d(\gamma)d\gamma. \tag{6}$$

If the delay $t_0 > 0$, then the above equation is rewritten as follows:

$$l(t) = e^{P(t-t_0)}l(t_0) + \int_{t_0}^t P(t-r)\varphi d(\gamma)d\gamma. \tag{7}$$

For for $t > t_0$, we have the following equation:

$$l(\gamma b + b) = \tau l(\gamma b) + \in d(\gamma b), \tag{8}$$

where $\tau = e^{Pb}$ and. $\varepsilon = \int_0^b e^{Pr}d\gamma\,\varphi$.
Applying equation (7) to (4), we obtain the following equation:

$$l((\gamma b + b)) = e^{Pb}l(\gamma b) + \int_{\gamma b}^{\gamma b + b} e^{P(\gamma b + b - \gamma)} d\gamma \varphi d(rb - T_{WT}) d\gamma,$$

$$= \tau l(\gamma b) + \int_{\gamma b}^{\gamma b + T_{WT}} e^{P(\gamma b + b - \gamma)} d\gamma \varphi d(rb - b) + \int_{\gamma b + T_{WT}}^{\gamma b + b} e^{P(\gamma b + b - \gamma)} d\gamma \varphi d(rb),$$

$$\gamma^1 = (\gamma b + b - \gamma), \tag{9}$$

$$\tau l(\gamma b) - \int_{b}^{b + T_{WT}} e^{P\gamma^1} d\gamma^1 \varphi d(\gamma b - b) - \int_{b - T_{WT}}^{0} e^{P\gamma^1} d\gamma^1 \varphi d(\gamma^1 b),$$

$$= \tau l(\gamma b) + \varepsilon_0 (T_{WT} d(rb)) + \varepsilon_1 (T_{WT} d(rb - b)).$$

In other scenarios, multiple copies of the signal are transmitted in the time interval and they take different routes while travelling corresponding to the direct path and scattered path. The spread in the delay indicated as σ_τ of 1 to 3 .... $h(\tau)$ is the delay profile. Taking the Fourier transform, we obtain $H(f) = \int_0^\infty h(\tau)e^{-J2\pi f\tau} d\tau$. The coherence bandwidth at which the delay profile response is almost flat, if the signal bandwidth $\beta_s < \beta_c$ is less than coherence of $\beta_c = 1/2\sigma_\tau$.

Figure 2 shows the signal for $\sigma_\tau \ll T_s$ or $\sigma_\tau \gg T_{signals}$. So, the sound signal interferes each other significantly and so on as delay spread increases to $\sigma_\tau > T_{signals}$ and $1/T_{signals} > 1/\sigma_\tau$, which implies to $\beta_s > 2\beta_c$ obtained as the interference. The T_x is stable and R_x is moving towards T_x, indicating the change in the frequency of sound varying due to relative motion between the T_x and R_x.

5.2. Simulation Results. In this section, we discuss the performance comparison of the proposed EELORP protocol by conducting simulations in Aqua-Sim [63–66]. Aqua-Sim is an extended version of NS-2 and offers easy implementation of underwater network scenarios. The parameters used for setting up the network are given in Table 2.

Using the simulations, we measure the energy consumption in nodes and the delay that occurred in the transmission of data in the UWSN. We also compare the results obtained by our proposed work with vector-based forwarding (VBF). Figure 3 shows the energy consumption by nodes in the network. From the results obtained, we can see that the nodes consume less energy using the proposed EELORP protocol compared to VBF protocol. Initially, the nodes have the same level of energy consumption with both the protocols, but as the number of nodes increases, the energy consumption using VBF becomes more compared to the proposed scheme. This signifies the better energy efficiency offered by the proposed protocol.

Figure 4 shows the delay incurred in transmission of the data packets using the protocols in the UWSN. From the results, we can see that using the proposed method EELORP and VBF, the delay incurred remains almost similar when the number of nodes is less. But as the number of nodes increases, the EELORP has less delay compared to VBF in the network. Thus, our results show that the proposed protocol can be used efficiently for numerous possibilities in underwater acoustic sensor networks with reduced delay.

6. Future Research Directions

6.1. Energy Efficiency. This has emerged as one of the major research areas for opportunistic routing protocols in underwater acoustic sensor networks. With restrictions and various limitations in recharging the sensor nodes, it is very important for any routing protocol to minimize the energy usage in the nodes while ensuring that the data gets delivered to the destination. Numerous protocols have tried to improve the energy efficiency in the network, but as UASN has an unpredictable nature, we should for further improvement in this research direction.

6.2. Channel Utilization. The unique features of UASNs like high propagation delay, constant mobility of sensor nodes, high error rate, and interference lead to a major challenge in ensuring the efficient utilization of the channel. Most of the existing protocols have various limitations in channel utilization and this area would be a major area of focus.

6.3. Communication Holes. Dealing with communication holes is a major challenge in UASNs, especially in networks with sparse deployment. Frequent movement of the sensor nodes due to currents and other reasons and failure of sensor nodes due to energy drainage or damage create void areas in the network. Thus, nodes will be unable to find suitable neighbor nodes to forward the data packet to the destination.

6.4. Security. Security of the data transmitted has become one of the major requirements of the applications deploying UASNs. It is therefore vital for all the routing protocols to include a security mechanism that can secure the data from any attackers.

6.5. Reliable Delivery. Reliable delivery of data packet at the destination is a major challenge in UASNs with a dynamic environment. Due to multiple reasons like damage of nodes, lack of energy, voids, etc., the data packet might get lost in the network. It is very important for any protocol to have strategies to manage any data loss and to make sure that the data reaches the destination, also keeping the number of retransmissions to a minimum to save the energy of nodes.
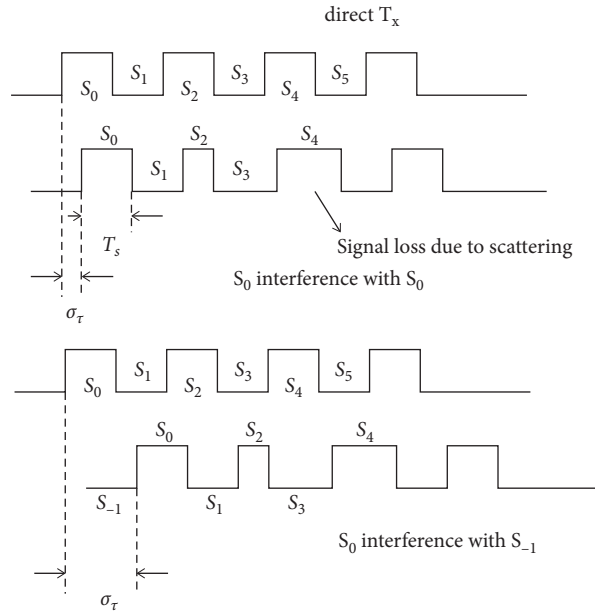
FIGURE 2: Physical layer-related losses and interference.

TABLE 2: Simulation specifications.

Parameter name	Values
Simulator name	NS 2.35 with aqua-sim
Dimension of topology	$1500 \times 1500 \times 1500$ m
Transmission range	250 m
Antenna type	Omni-directional
Data rate	50 kbps
Packet size	25 to 125 bytes
Number of nodes	100 to 300
Simulation time	200 s
Number of simulation runs	10
Protocols	EELOP, VBF
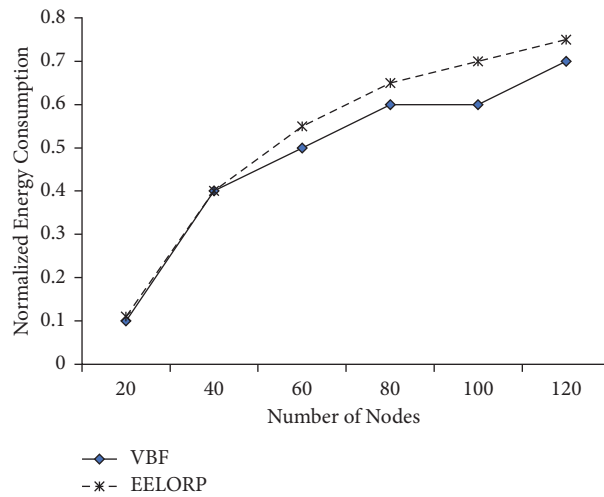
FIGURE 3: Normalized energy consumption versus the number of nodes.
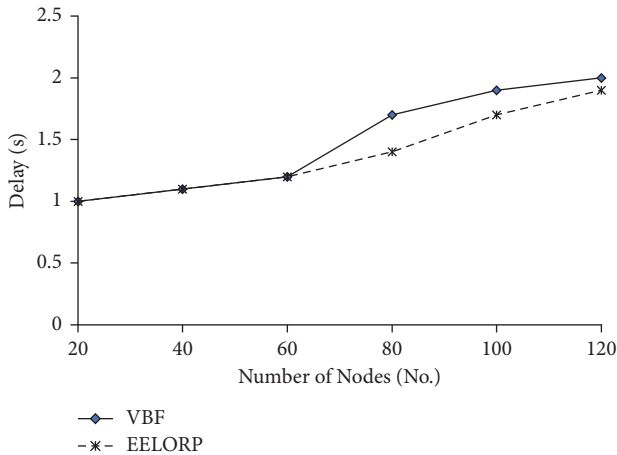
FIGURE 4: Delay versus number of nodes.

7. Conclusion

This study presented a systematic survey on the location-based opportunistic routing protocols in underwater acoustic sensor networks. The study initially discussed the working of underwater sensor networks, the challenges and issues, the latest applications using UASNs, and the working of opportunistic routing in underwater acoustic sensor networks. A detailed discussion on all recently proposed location-based opportunistic routing protocols was presented with a focus on their design and working. A discussion on the design and working of an easy-to-implement energy-efficient location-based opportunistic routing protocol (EELORP) that can be used efficiently for numerous possibilities in underwater acoustic sensor networks with reduced delay was presented. A discussion on results obtained with simulations was then presented along with comparisons with existing protocols. Finally, a brief discussion on the future research directions was presented.

Data Availability

All the output data with details can be accessed through the first author at reasonable request (email: varunmenon@ieee.org).

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] T. Qiu, Z. Zhao, T. Zhang, C. Chen, and C. L. P. Chen, "Underwater Internet of things in smart ocean: system Architecture and open issues," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4297–4307, 2000.

[2] G. Han, C. Zhang, L. Shu, and J. J. P. C. Rodrigues, "Impacts of deployment strategies on localization performance in underwater acoustic sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1725–1733, 2015.

[3] S. Fattah, A. Gani, I. Ahmedy, M. Y. I. Idris, and I. A. Targio Hashem, "A survey on underwater wireless sensor networks: requirements, taxonomy, recent advances, and open research challenges," *Sensors*, vol. 20, no. 18, p. 5393, 2020.

[4] G. Han, X. Long, C. Zhu, M. Guizani, and W. Zhang, "A high-availability data collection scheme based on multi-AUVs for underwater sensor networks," *IEEE Transactions on Mobile Computing*, vol. 19, no. 5, pp. 1010–1022, 2020.

[5] A. M. Khasawneh, O. Kaiwartya, A. Khalifeh, L. M. Abualigah, and J. Lloret, "Green computing in underwater wireless sensor networks pressure centric energy modeling," *IEEE Systems Journal*, vol. 14, no. 4, pp. 4735–4745, 2020.

[6] V. Menon, D. Midhunchakkaravarthy, S. Verma, A. Sujith, and M. Manju, "Enabling reliable communication in Internet of underwater things: applications, challenges and future directions," in *Proceedings of the 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*, pp. 296–301, Jalandhar, India, May 2021.

[7] G. Han, H. Wang, J. A. Ansere, J. Jiang, and Y. Peng, "SSLP: a stratification-based source location privacy scheme in underwater acoustic sensor networks," *IEEE Network*, vol. 34, no. 4, pp. 188–195, 2020.

[8] V. G. Menon, "Opportunistic routing protocols in underwater acoustic sensor networks: issues, challenges, and future directions," *Magnetic Communications*, CRC Press, FL, USA, pp. 109–130, 2018.

[9] H. Cui, Y. Zhang, X. Liu, and D. Sun, "The simulation and emulation platforms of underwater acoustic sensor networks," in *Proceedings of the IEEE/OES China Ocean Acoustics (COA)*, pp. 1–5, Harbin, China, January 2016.

[10] Y. Chen, X. Jin, and X. Xu, "Mobile data collection paths for node cooperative underwater acoustic sensor networks," in *Proceedings of the OCEANS*, pp. 1–5, Shanghai,China, April 2016.

[11] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, "A survey of architectures and localization techniques for underwater acoustic sensor networks," *IEEE Communications Surveys & Tutorials, in IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 487–502, 2011.

[12] K. S. Keerthi, B. Mahapatra, and V. G. Menon, "Into the world of underwater swarm robotics: architecture, communication, applications and challenges," *Recent Advances in Computer Science and Communications*, vol. 13, no. 2, pp. 110–119, 2020.

[13] J. Yan, Y. Gong, C. Chen, X. Luo, and X. Guan, "AUV-aided localization for Internet of underwater things: a reinforcement-learning-based method," *IEEE Internet of Things JournalEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9728–9746, 2020.

[14] W. Zhang, G. Han, X. Wang, M. Guizani, K. Fan, and L. Shu, "A node location algorithm based on node movement prediction in underwater acoustic sensor networks," *IEEE Transactions on Vehicular Technology, IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3166–3178, 2020.

[15] Y. Song, "Underwater acoustic sensor networks with cost efficiency for Internet of underwater things," *IEEE Transactions on Industrial Electronics, in IEEE Transactions on Industrial Electronics*, vol. 68, no. 2, pp. 1707–1716, 2021.

[16] S. Zheng, X. Cao, F. Tong, G. Zhang, and Y. Dong, "Performance evaluation of acoustic network for underwater autonomous vehicle in confined spaces," in *Proceedings of the IEEE 8th International Conference on Underwater System Technology: Theory and Applications (USYS)*, pp. 1–4, Wuhan, China, December 2018.

[17] S. Y. Kulik, A. Y. Rodionov, F. S. Dubrovin, and P. P. Unru, "On reliability of data transmission and distance estimation using mobile underwater acoustic modems," in *Proceedings of*

the 25th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), pp. 1–4, St. Petersburg, Russia, May 2018.

[18] S. Biswas and R. Morris, "ExOR, ACM SIGCOMM - Computer Communication Review," in *Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 133–144, USA, October 2005.

[19] M. Ismail, M. Islam, I. Ahmad, F. Aslam, and A. Baseer, "Reliable path selection and opportunistic routing protocol for underwater wireless sensor networks," *IEEE Access*, vol. 8, pp. 100346–100364, 2020.

[20] V. G. Menon and P. M. J. Prathap, "Comparative analysis of opportunistic routing protocols for underwater acoustic sensor networks," in *Proceedings of the International Conference on Emerging Technological Trends (ICETT)*, pp. 1–5, Kollam, India, October 2016.

[21] P. Xie, J. Cui, and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks," *NETWORKING 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, vol. 3976, pp. 1216–1221, 2006.

[22] H. Daeyoup and D. Kim, "DFR: directional flooding-based routing protocol for underwater sensor networks," in *Proceedings of the IEEE Oceans*, Quebec, QC, Canada, September 2007.

[23] W. Liang, H. Yu, L. Liu, B. Li, and C. Che, "Information-carrying based routing protocol for underwater acoustic sensor network," in *Proceedings of the International Conference on Mechatronics and Automation (ICMA 2007)*, pp. 729–734, Harbin, China, August 2007.

[24] N. Nicolaou, A. See, P. Xie, J. H Cui, and D. Maggiorini, "Improving the robustness of location-based routing for underwater sensor networks," in *In Proceedings of the IEEE Oceans*, Aberdeen, UK, September 2008.

[25] J. Chen, X. Wu, and G. Rebar Chen, "A reliable and energy balanced routing algorithm for UWSNs," in *Proceedings of the 7th International Conference on Grid and Cooperative Computing (GCC'08)*, pp. 349–355, Shenzhen, China, October 2008.

[26] P. Xie, Z. Zhou, Z. Peng, J. –H. Cui, and Z. Chi, "Void Avoidance in three-dimensional mobile underwater sensor networks," in *In Proceedings of the the 4th International Conference on Wireless Algorithms, Systems and Applications*, Boston, MA, USA, August 2009.

[27] M. V. Priya, A. A. Kumari, Traffic aware multipath communication for time-critical applications in underwater acoustic sensor networks," *International Journal of Managment, IT and Engineering*, vol. 2, pp. 66–73, 2012.

[28] P. Wang, D.-hao Fu, C.-qing Zhao, J.-chun Xing, Qi-liang Yang, and X.-fei Du, "A reliable and efficient routing protocol for Underwater Acoustic Sensor Networks," in *Proceedings of the IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*, Nanjing, China, May 2013.

[29] J. She, J. Wang, J. Wang, J. Zhang, and S. Wang, "Location-aware routing protocol for underwater sensor networks," in *Proceedings of the Advanced Technologies, Embedded and Multimedia for Human-Centric Computing*, November 2014.

[30] J. Park, S. Lee, D. Kim, and Y. Hong, "QoS-aware directional flooding-based routing for underwater wireless sensor networks," in *Proceedings of the WUWNET 14th International*

[31] I. Naveed, A. TurkiAli, N. F. Muhammad et al., "SEDG: Scalable and Efficient Data Gathering Routing Protocol for Underwater WSNs," *Procedia Computer Science*, vol. 2, pp. 568–575, Elsevier, Amsterdam, Netherland, 2015.

[32] I. Naveed, A. Turki, N. F. Muhammad et al., "AEDG: AUV-aided Efficient Data Gathering Routing Protocol for Underwater Wireless Sensor Network," *Elsevier, Procedia Computer Science*, vol. 2, pp. 568–575, 2015.

[33] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "Modeling and analysis of opportunistic routing in low duty-cycle underwater sensor networks," in *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 125–132, New York, NY, USA, November 2015.

[34] S. Basagni, C. Petrioli, R. Petroccia, and D. S.. Carp, "A channel-aware routing protocol for underwater acoustic wireless networks," *Ad Hoc Networks*, vol. 34, pp. 92–104, 2015.

[35] Q. Wang, C. Fei, L. Zhi, and Q. Qian, "A novel efficient forwarding protocol for 3-D underwater wireless sensor networks," in *In Proceedings of the IEEE 11th International Conference on Industrial Electronics and Applications*, Hefei, China, June 2016.

[36] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "Geographic and opportunistic routing for under water sensor networks," *IEEE Transactions on Computers*, vol. 65, pp. 548–561, 2016.

[37] D. Li, J. Du, and L. Liu, "A data routing algorithm based on Markov model in underwater wireless sensor networks," in *Proceedings of the IEEE 16th International Conference on Ubiquitous Wireless Broadband*, Nanjing, China, December 2016.

[38] S. Han, Y. Yue, W. Meng, and X. Wu, "A localization-based routing protocol for dynamic underwater sensor networks," in *Proceedings of the IEEE Global Commuications Conference*, Washington, DC, USA, December 2016.

[39] N. Kanthimathi and Dejey, "Void handling using Geo-Opportunistic Routing in underwater wireless sensor networks," *Computers & Electrical Engineering*, vol. 64, pp. 365–379, 2017.

[40] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "Performance modeling and analysis of void-handling methodologies in underwater wireless sensor networks," *Computer Networks*, vol. 126, pp. 1–14, 2017.

[41] M. Yusuf and S. Uddin, "Low-overhead range-based 3D localization technique for underwater sensor networks," in *Proceedings of the IEEE International Conference on Communications*, Kuala Lumpur, Malaysia, May 2016.

[42] E. Isufi, H. Dol, and G. Leus, "Advanced flooding-based routing protocols for underwater sensor networks," *EURASIP Journal on Applied Signal Processing*, vol. 2016, pp. 1–12, 2016.

[43] K. Wang, H. Gao, X. Xu, J. Jiag, and D. Yue, "An energy-efficient reliable data transmission scheme for complex environmental monitoring in underwater acoustic sensor networks," *IEEE Sensors Journal*, vol. 16, pp. 4051–4062, 2016.

[44] A. walayat, N. Javaid, M. Akbar, and Z. A. Khan, "MEES: mobile energy efficient square routing for underwater wireless sensor networks," in *Proceedings of the IEEE 31st International Conference on Advanced Information Networking and Applications*, Taipei, Taiwan, May 2017.

[45] I. Yazgi and B. Baykal, "Topology control vector based forwarding algorithm for underwater acoustic networks," in

Conference on Underwater Networks and Systems, Rome, Italy, November 2014.

Proceedings of the IEEE 24th International Conference on Signal Processing and Communication Application, Zonguldak, Turkey, May2016.

[46] F. A. Salti, N. Alzeidi, and B. R. Arafeh, "EMGGR: an energy-efficient multipath grid-based geographic routing protocol for underwater wireless sensor networks," *Wireless Networks*, vol. 23, no. 4, pp. 1301–1314, 2017.

[47] N. Kanthimathi and Dejey, "Balanced and Multi-objective Optimized Opportunistic Routing for Underwater Sensor Networks," *Wireless Personal Communications*, vol. 94, no. 4, pp. 2417–2440, 2017.

[48] A. Khan, N. Javaid, I. Ali, and H. A. Mohammad, "An energy efficient interference aware routing protocol for underwater WSNs," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 11 10, 2017.

[49] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Coutinho, "PCR: a power control-based opportunistic routing for underwater sensor networks," in *Proceedings of the MSWiM 2018 - Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 173–180, New York, NY, USA, October 2018.

[50] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "A stateless opportunistic routing protocol for underwater sensor networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 8237351, 18 pages, 2018.

[51] Z. Jin, N. Wang, Y. Su, and Q. Yang, "A glider-assisted link disruption restoration mechanism in underwater acoustic sensor networks," *Sensors*, vol. 18, no. 2, 2018.

[52] W. Zhuo, H. Guangjie, Q. Hongde, Z. Suping, and S. Yancheng, "An energy - aware and void-avoidable routing protocol for underwater sensor networks," *IEEE Access, Journals & Magazines*, vol. 6, pp. 7792–7801, 2018.

[53] R. Bu, S. Wang, and H. Wang, "Fuzzy logic vector–based forwarding routing protocol for underwater acoustic sensor networks," *Trans. Emerg. Telecommun. Technol.*vol. 29, pp. 1–18, 2018.

[54] F. Ahmed, Z. Wadud, N. Javaid, N. Alrajeh, M. S. Alabed, and U. Qasim, "Mobile sinks assisted geographic and opportunistic routing-based interference avoidance for underwater wireless sensor network," *Sensors*, vol. 18, no. 4, 2018.

[55] Z. Rahman, F. Hashim, M. F. A. Rasid, and M. Othman, "Totally opportunistic routing algorithm (TORA) for underwater wireless sensor network," *PLoS One*, vol. 13, no. 6, 2018.

[56] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "EnOR: Energy balancing routing protocol for underwater sensor networks," in *Proceedings of the IEEE International Conference on Communications*, Paris,France, May 2017.

[57] I. U. Khan, M. Islam, M. Ismail et al., "Adaptive hop-by-hop cone vector-based forwarding protocol for underwater wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 16, no. 9, Article ID 15501477209, 2020.

[58] M. Zhang and W. Cai, "Energy-efficient depth based probabilistic routing within 2-hop neighborhood for underwater sensor networks," *in IEEE Sensors Letters*, vol. 4, no. 6, pp. 1–4, 2020.

[59] Y. Chen, J. Zhu, L. Wan, S. Huang, X. Zhang, and X. Xu, "ACOA-AFSA fusion dynamic coded cooperation routing for different scale multi-hop underwater acoustic sensor networks," *in IEEE Access*, vol. 8, Article ID 186788, 2020.

[60] Y. Chen, K. Zheng, X. Fang, L. Wan, and X. Xu, "QMCR: a Q-learning-based multi-hop cooperative routing protocol for underwater acoustic sensor networks," *in China Communications*, vol. 18, no. 8, pp. 224–236, 2021.

[61] D. Zhao, G. Lun, and R. Xue, "Coding-aware opportunistic routing for sparse underwater wireless sensor networks," *IEEE Access*, vol. 9, Article ID 50187, 2021.

[62] Z. Jin, Q. Zhao, and Y. Luo, "Routing void prediction and repairing in AUV-assisted underwater acoustic sensor networks," *IEEE Access*, vol. 8, Article ID 54212, 2020.

[63] M. R. Khosravi, H. Basri, and H. Rostami, "Efficient routing for dense UWSNs with high-speed mobile nodes using spherical divisions," *The Journal of Supercomputing*, vol. 74, no. 2, pp. 696–716, 2018.

[64] M. R. Khosravi, H. Basri, H. Rostami, and S. Samadi, "Distributed random cooperation for VBF-based routing in high-speed dense underwater acoustic sensor networks," *The Journal of Supercomputing*, vol. 74, no. 11, pp. 6184–6200, 2018.

[65] M. R. Khosravi, "The shortfalls of underwater sensor network simulators," *Sea Technology*, vol. 60, no. 5, p. 41, 2019.

[66] P. Xie, Z. Zhou, Z. Peng et al., "Aqua-Sim: an NS-2 based simulator for underwater sensor networks," in *Proceedings of the Oceans 2009*, pp. 1–7, Biloxi, MS, USA, October 2009.

Research Article

Real-Time Modulation of Physical Training Intensity Based on Wavelet Recursive Fuzzy Neural Networks

Wenzhou Fang [iD],[1] Lili Wang [iD],[1] Xinxin Liao [iD],[2] and Miao Tan [iD][1]

[1]*Shenyang Sport University, Shenyang, Liaoning 110102, China*
[2]*University Putra Malaysia, UPM Aerdang, Selangor Darul Ehsan 43400, Malaysia*

Correspondence should be addressed to Miao Tan; tangm@syty.edu.cn

In this study, a wavelet recurrent fuzzy neural network is used to conduct in-depth research and analysis on the real-time regulation of physical training intensity. Firstly, an inter-process control technique is proposed to solve the problem of incomplete control flow graph construction caused by the inability to effectively collect all program control flow information in the process of static analysis, in preparation for the research of fuzzy testing technique. Next, a wavelet recursive fuzzy neural network-guided fuzzy testing technique is proposed to solve the problem of fuzzy tests falling into invalid variation due to the lack of directionality in the fuzzy testing process. Each neuron in the feedforward network is divided into different groups according to the order of receiving information. Each group can be regarded as a neural layer. The neurons in each layer receive the output of the neurons in the previous layer and output to the neurons in the next layer. The empirical data show that injury-preventive fitness training can effectively improve all physical qualities in the first phase of preparation and can effectively maintain the physical state and effectively contribute to their abilities during the competition period, and its injury-preventive fitness training interventions were verified by statistical analysis to have a dangerous main effect on their pre and post-test performance. Therefore, it is still not possible to determine its correlation with the coordination and improvement of the athletes' physical fitness, and the integration of the basic physical training and rehabilitation physical training systems, making this theory a new special training theory.

1. Introduction

Physical fitness is based on the energy metabolic activities of the three major energy supply systems of the human body and is an important component of athletes' competitive ability, as expressed through the skeletal muscle system. Physical fitness training is one of the main ways to improve the athletes' special athletic ability and special performance. With the continuous updating of world records in competitive sports, the important role of physical training is becoming increasingly significant. Based on the data model of each training stage, the stage goal is set and the training goal is "digitized." The data analysis and feedback method is then used to represent and evaluate the achievement of the athletes' goals at each stage of training, and if the desired numerical goals are not achieved, the reasons for this are sought and adjustments are made until the goals are achieved [1]. Sports injuries are a topic that every athlete faces and are the most likely to occur during training and competition. The factors that occur are related to the athlete's type of sport, technique, and tactics, basic physical athletic ability, training and game facility environment, and psychological factors. Basketball is also one of the sports with a high injury rate due to high-intensity confrontation and high-speed rushing, under excessive use of joints muscles and ligaments [2]. Modern basketball is very powerful in terms of physical collision and confrontation, so the risk of injury to athletes during training and competition increases. The nuisance of injuries throughout different intensity levels of basketball has been a pressing issue. It is the only interface between genetic algorithm and specific application problems and is the only basis for natural selection. The genetic operations that change the internal structure of the population are all controlled by the evaluation function. Many types of

evaluation functions can be used in genetic algorithms, but they should satisfy the following condition: the function values are partially ordered. From studies and reports, it has been found that common acute injuries among basketball players include bruises, contusions, fractures, joint sprains, and muscle strains. The incidence of sports injuries is as high as 71.1%. Sports injuries not only hurt the physical and mental health of individual athletes but also affect their competitive level and restrict the development of the whole team, forming a vicious circle [3].

Injury prevention physical training is a combination of rehabilitation physical training, basic physical training, special physical training, and special technical and tactical needs together to form ways and methods, applied to athletes in different special physical training sports, injury, and functional state supervision to reduce the risk of damage and injury of athletes in the competitive state. It combines the functions of physical training and rehabilitation training [4]. Physical training as an intervention is supported by knowledge of rehabilitation medicine to ensure physical fitness for daily training. The main purpose is to improve and develop the quality of training by providing a basis for recovery training after an athlete's injury and by combining it with an injury prevention-based training philosophy.

Therefore, the inter-test path feedback is needed to use the insert pile technique to use the algorithm to guide the sample variation at the right time to achieve the predefined fuzzy test effect. The stubbing technique is carried out by adding a portion of code to the front of each basic block of the program, which gets executed during the execution of the program [5]. The purpose of this study is to investigate the digital control of physical training of athletes. By creating a relevant physical training monitoring and evaluation system, thus representing the entire training process of the athlete with data. This includes the development of the plan, the statistics of the training data, the evaluation of the stage training effect, and the feedback of the stage training effect. During training, the activation function saturates due to the excessive adjustment of the weights, so that the adjustment of the network weights almost stagnates. To avoid this situation, one is to choose a smaller initial weight, and the other is to use a smaller learning rate. The BP algorithm can make the network weights converge to a final solution, but it cannot guarantee the global optimal solution of the error hyperplane, or it may be a local minimum. The digital control is truly achieved so that the athletes' fitness level develops in the expected direction, which in turn provides a digital basis for athletes' fitness training and makes a slight contribution to promoting the development of the sport.

In the second part of this study, the existing research is analysed and explained and the shortcomings of the existing research are presented. In the third part of this study, we focus on the design of wavelet recursive fuzzy neural network algorithm and the real-time regulation system of physical training intensity. In the fourth part, we provide a detailed analysis of the neural network algorithm and the performance results of the system. In the last part of this study, we summarize and explain the results of this study.

2. Related Works

The influencing factors and implementation effects are comprehensively analyzed. The training process is arranged in slightly different proportions, with the barbell training method as the main body and comprehensive equipment as the auxiliary training. The method of administration required for basketball-specific strength is experimentally corrected, verified that a variety of combined training methods can reduce the excessive fatigue and sports injuries of basketball players [6]. It was pointed out that their characteristics and the choice of basketball training intensity methods must be compatible with each other. The development of basic and special strength training is directed towards training in centrifugal and centripetal, positive, and negative forces, static to dynamic balance, power chains, and responsiveness, and the generation of power differs from the traditional speed and explosive power in muscle power, which is produced by neuromodulation of attack energy [7]. It introduces the concept of strength quality, illustration of body strength training in different parts, classification and description of characteristics according to body position and movement characteristics, intensity division, and flexibility training principles and methods to develop a good basis for training methods and overall training rules, which can be used as a basis and direction for the promotion of special strength of each sport; there are overall strength training methods according to the body power chain and training patterns [8]. With the rehabilitation and injury prevention training, we hope to prevent the incidence of sports injuries and have the effect of postinjury rehabilitation [9].

Athletic training is a scientific and theoretical system of great complexity consisting of multiple disciplines and the cross-organization of multiple training theories and training methods guided by the corresponding theories [10]. This object maintains the trend of motion in its direction of motion. The mechanical motion of the object does not occur in isolation, it interacts with the surrounding objects, and this interaction is manifested as the mechanical motion of the moving object and the surrounding objects in the transfer process. Momentum is a physical quantity that measures mechanical motion from the perspective of mechanical motion transmission. The overall system of physical training must analyse evaluating the actual situation. The overall planning, implementation, operation of the training cycle, body load monitoring, nutrition and rehabilitation, injury prevention and fatigue recovery, and evaluation and monitoring of training are composed of seven components, which are different subsystems [11]. The assessment of explosive power and speed qualities refers to the ability of the athlete to reach the fastest speed shown in the shortest possible time with and without changing the direction of travel, and the athlete is tested based on maximum heart rate and respiratory rate for an appropriate period; the assessment of agility qualities refers to the athlete's ability to decelerate to 0 and accelerate again during a change in direction of travel [12].

The motion of a sliding mode variable structure control can be divided into two processes, namely, converging mode

and sliding mode. First, the control force will make the system state converge to the sliding mode surface motion before the system does not move to the sliding mode surface, and when the system reaches the sliding mode surface, the control law on the sliding mode surface is determined by the selected sliding mode surface and is no longer affected by external disturbances, which can ensure the stable operation of the system [13].

2.1. Wavelet Recursive Fuzzy Neural Network Physical Training Intensity Real-Time Modulation Analysis

2.1.1. Wavelet Recursive Fuzzy Neural Network Algorithm Design. Neural network control is an artificial intelligence system that emulates the human thought process; however, the initial neurons in neural networks had only two states, on and off, which could not reflect the highly nonlinear characteristics of the actual neurons [14]. As researchers in various fields at home and abroad successively invested in research, it created the rise of artificial intelligence control methods.

The input signal $x_1, x_2, \ldots, x_n$ to the basic structure of a neuron is the input signal from other neurons to this neuron. The signal path is like that of the input nerve on the actual neuron structure, that is, the dendrites, which transmit and receive information from other nerve cells. The input signal to the artificial neuron is adjusted by connection weights $w_1, w_2, \ldots, w_n$, which can be positive or negative, with positive connection weights indicating growth and negative connection weights indicating negative inhibition. Aiming at this problem, a method is proposed to minimize the conflict between the convergence speed and the steady-state error by introducing the momentum term and optimizing the momentum factor. The algorithm reduces the cost function of nonlinear principal component analysis the fastest through the optimized momentum term, so that the convergence speed of the algorithm is accelerated. After the input signals weighted by the connection weights are all summer, the excitation function on the artificial neuron is nonlinearly transformed to obtain a new output signal, which is then communicated to the other neurons of the system, and its function is equivalent to the nucleus of a nerve cell on the actual neuron structure:

$$\text{net}_j = \sum_{i=1}^{n} \left(w_{ji}^2 x_i \right) + b_{ij},$$
$$y_j = f\left(\text{net}_j\right) + b_j, \tag{1}$$

where w_{ji}^2 shows the value of the connection weight of the i-th input to the j-th neuron, b_j denotes the bias of the j-th neuron, $f(\text{net}_j)$ is the nonlinear excitation function of the neuron and converts the sum of the product weights at the input to the value at the output. A single neuron can process and learn data if multiple artificial neurons can be combined into a complete neural network after some proper arrangement. The number of layers of the designed neural

network can be decided according to the complexity of the problem to be solved, the classification of neural networks in terms of the number of layers can be broadly classified as single-layer neural networks and multi-layer neural networks, and its structure is shown in Figure 1.

As the neural network can learn online, the fuzzy system has the characteristics of logical description and judgment, wavelet processing can analyse time-varying signals, and the recursive structure gives the system better dynamic characteristics [15]. The recursive wavelet fuzzy neural network is structurally divided into a total of five layers, with three implicit layers, namely, the affiliation function layer, the rule layer, and the recursive wavelet function layer. The neural network controller combines the neural network with fuzzy logic, wavelet processing, and recursive structure to improve its processing capability and accuracy and to solve the shortcomings of static mapping. The transfer relationship between each layer will be explained next.

Both neuron nodes in this layer are input nodes, which are equivalent to the input variables. The error of the H-type motion platform includes the position synchronization error of the two axes and the speed synchronization error of the two axes. The linear transformation relationship between the input and output of the neurons in this layer can be expressed as follows:

$$\text{net}_j^1 (N) = x_i^2,$$
$$y_j^1 (N) = \text{net}_j^1 (N), \quad j = 1, 2, \tag{2}$$

where x_i^2 is the input signal of the input layer, and the input variables are the position synchronization error and velocity synchronization error $x_1^1 = e_{y_1} + e_{y_2}$, e_{y_1} and the e_{y_2} position tracking error of the Y_1 and Y_2 linear motors, respectively; $y^1 (N)$ is the output signal of the input layer; N is the number of samples.

$$y_j^2 (N) = f_j^2 \left(exp \, \text{net}_j^2 (N) N \right), \quad j = 1, 2, \ldots, 6. \tag{3}$$

The output of each neuron in the input layer corresponds to 3 neurons in the affiliation function layer, the nonlinear transformation in the affiliation function layer uses Gaussian function, and this transformation method incorporates the fuzzy logic inference method so that the inductive performance of the network is improved. The mutation direction is provided for the sample mutation; then the neural network is used to build a program flow flattening model, the branch logic of the program is simulated smoothly, and a more effective mutation position is obtained through gradient calculation. Here, $x_i^2 (N) = y_i^2 (N)$ is the output of the input layer; m_j is the mean of the Gaussian function of the affiliation function layer; σ_j is the standard deviation of the Gaussian function of the affiliation function layer; $y_i^2 (N)$ is the output of the neurons of the affiliation function layer. This layer contains wavelet function operations, recursive operations, and the posterior part of fuzzy logic rules. The output of the wavelet function is ψ_k, denoted as follows:
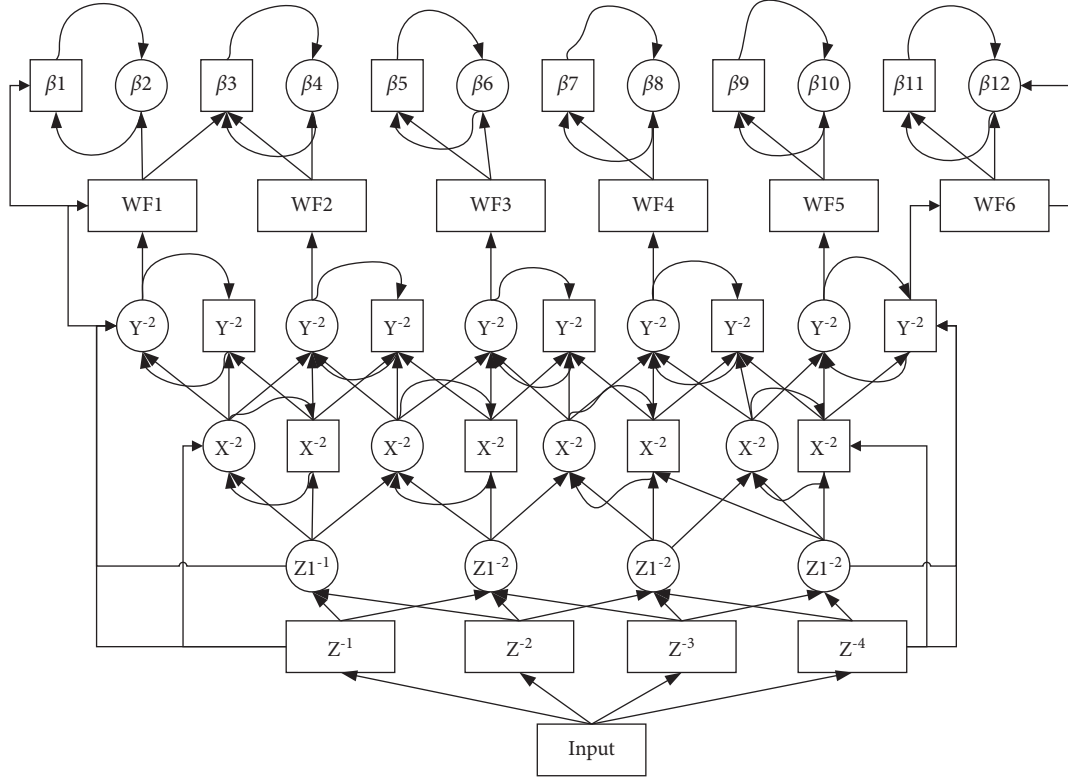
FIGURE 1: Wavelet recurrent type neural network.

$$\phi_{ik}\left(x_i^1\right) = \frac{1}{\sqrt{|b_{ik}|}}\left[1 + \frac{\left(x_i^1 - a_{ik}^2\right)}{b_{ik}^2}\right]\exp\left[1 - \frac{\left(x_i^1 - a_{ik}^2\right)^2}{2b_{ik}^2}\right],$$
(4)

$$\psi_k = \sum_i w_{ik}^2 \phi_{ik}\left(x_i^1\right).$$

where is the i-th wavelet function in the k-th neuron of ϕ_{ik} this layer; ψ_k is the output of the k-th wavelet function; x_i^1 is the connection weight of the wavelet function; a_{ik} and b_{ik} are the translation factor and scaling factor of the wavelet function, respectively. From Figure 2, we can see the relationship between different translation factors and scaling factors on the input and output of wavelet functions.

The gradient descent method, also known as the error backpropagation algorithm, is a basic learning algorithm in the regulation process of neural network systems. The computational process of the gradient descent method is to solve for the minimum value along the direction of the decreasing gradient or the maximum value along the direction of the increasing gradient and to adjust the parameters in the neural network in a feedback manner so that the system error gradually approaches the minimum value. The energy function V is first defined as follows:

$$V = \frac{1}{4}e^2.$$
(5)

The recursive wavelet fuzzy neural network with parameter learning updates the iterative algorithm in the output layer with the error term, as shown in the following equation:
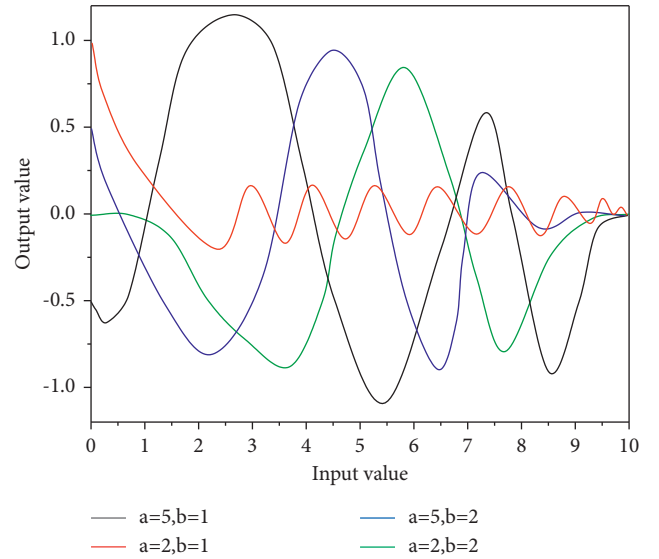
FIGURE 2: Schematic diagram of the wavelet function.

$$\delta_o^5 = \frac{\partial V}{\partial e}\frac{\partial e}{\partial y_o^5}.$$
(6)

The algorithm first establishes the most ideal access network and the least ideal access network. The attribute parameters of the most ideal access network are relatively good, and a positive ideal solution can be obtained by normalizing the attribute parameters of the most ideal

access network. However, the attribute parameters of the least ideal access network are relatively poor, and a negative ideal solution can be obtained by normalizing the attribute parameters of the least ideal access network. By calculating the geometric distance between each candidate network and the positive and negative ideal solutions, the proximity of each candidate network to the positive and negative ideal solutions is obtained. These closenesses are then sorted. The closest network is the one closest to the ideal solution [16].

The fourth layer of the fuzzy neural network model is the fuzzy inference layer. Since each alternative network score corresponds to three layers (L, M, and H), there are a total of 3 M nodes in this layer. The input and output of performing the operation at the fourth layer node yield are expressed as

$$I_i^4 = \sum_{j \in G_i} x_{ij}^4,$$
$$O_i^4 = \max\left(1, I_i^4\right), \tag{7}$$

The first part fuzzifies the input alternative network parameters, establishes fuzzy rules, outputs the scores of the alternative networks, and relies on the adaptive learning capability of the neural network for error backpropagation in the process of network selection to adaptively adjust the affiliation function parameters, which in turn enables the fuzzy neural network to be trained. After the scores of each alternative network are obtained through training, the scores of each alternative network are weighted according to the degree of preference of different business users for each alternative network, and the alternative network with the highest weighted score is selected to serve the users. The reason is found in time and the corresponding adjustments are made until the goal is achieved. Sports injury is a topic that every athlete will face, and it is the most likely to occur in training and competition. The main part of the second part is to use the dragonfly algorithm to find the optimal initial parameters of the Gaussian affiliation function of the fuzzy neural network when no fuzzy neural network training is performed until the end of the iteration, and the food location found by the dragonfly population is the optimal initial parameter of the Gaussian affiliation function.

2.1.2. Physical Training Intensity Real-Time Regulation System Design. The motion data acquisition side consists of a motion acquisition module and an adjustable weighted dumbbell. The motion acquisition module is placed in a 3D printed circular housing with dimensions of 104 mm outer diameter, 20 mm inner diameter, and 20 mm thickness, which is fixed directly on the threaded rod. The motion acquisition module mainly includes a microprocessor module, a power supply module, and a sensor module. Among them, the microprocessor module uses an ESP8266 microcontroller with a built-in Wi-Fi module, the sensor module uses a JY901 motion sensor,

and the power supply module uses a lithium polymer battery and charging circuit [17]. The microprocessor module is the core of the system, and the ESP8266 developed by Loxin is used in this system, which provides a highly integrated Wi-Fi SoC solution with low power consumption, compact design, and high stability to meet the requirements of the system.

The ESP8266 has a built-in ultra-low-power Ten silica L106 32 bit RISC processor with a CPU clock speed of up to 160 MHz and supports real-time operating systems (RTOSs) and Wi-Fi protocol stacks, leaving up to 80% of the processing power for application programming and development. Its high level of integration, with standard digital peripheral interfaces, antenna switches, RF, power amplifiers, low-noise amplifiers, filters, and power management modules, allows it to be significantly reduced in size. At the same time, the ESP8266 features low power consumption and is designed for mobile devices, wearable electronics, and IoT applications, while having ultra-low power consumption. The functional principle of the ESP8266 is shown in Figure 3.

After powering up the data acquisition module, the program first initializes the MCU and the sensor module, searches for a Wi-Fi network, and joins it to establish a secondary peer-to-peer network. At this point, the MCU starts collecting data from the sensor module. A brief description of the process of the distribution network follows. There are two cases in which the ESP8266 needs to be wired to connect to a specified Wi-Fi network. The first is when the ESP8266 is first powered up, and the second is when the user has changed the network environment. The traditional method of network pairing is to write the wireless network to be connected into the ESP8266 program, and when the module is powered on, it automatically searches for the specified wireless network and connects to it. This pairing method is easy and fast but has the drawback that it cannot be changed when the wireless network environment changes or when the user needs to reconnect to another wireless network. Therefore, there is a need to improve the network allocation procedure to be able to achieve more flexible network allocation.

In the data layer, the server receives the data information from the collector and the client and then needs to store the data on the server and perform some related data processing. Therefore, a database of user information and user workout information needs to be created for easy management. The server uses an SQLite database, and the established database contains the following data tables as listed in Table 1.

The collected acceleration signal line is analysed to determine whether the signal needs low-pass or high-pass filtering processing. After a lot of experiments, it was found that the average time to do a dumbbell movement is about 3 s, which corresponds to a frequency of 1/3 Hz, which belongs to the low-frequency part of the signal [18]. It is used under the supervision of athletes in different special physical training sports, injuries, and functional states to reduce the risk of injury and injury of athletes in
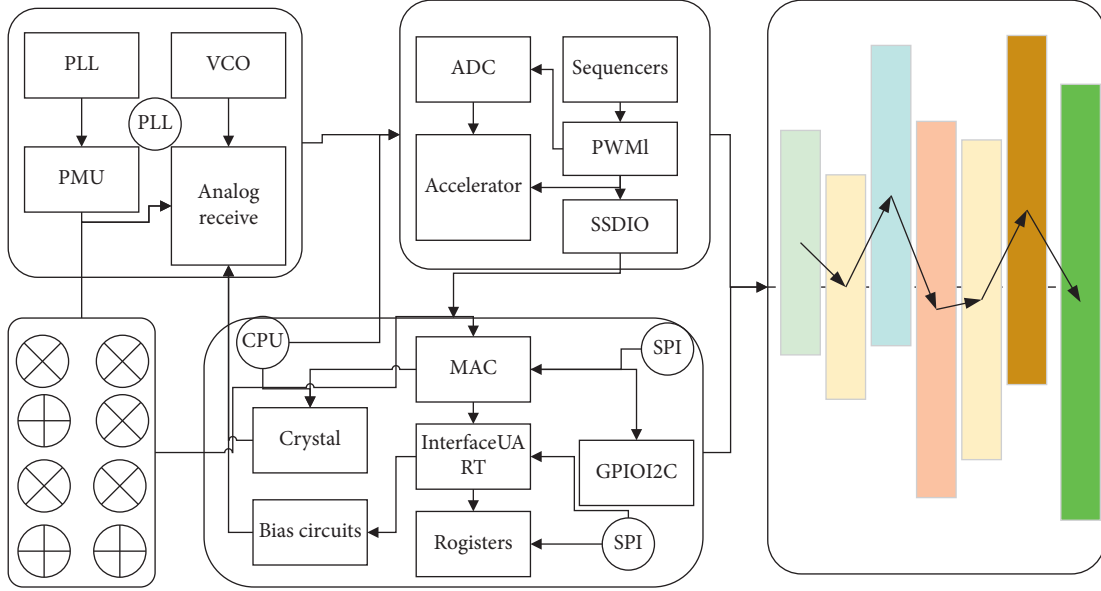
FIGURE 3: Functional schematic.

TABLE 1: Real-time motion data table.

Field name	Type	Illustratation
_id	double	Y-axis acceleration
xAcc	double	Z-axis acceleration
Field name	double	X-axis angular velocity
yAcc	double	Y-axis angular velocity
zAcc	double	Z-axis angular velocity
xPalstance	double	X-axis angle
yPalstance	double	Y-axis angle

competitive conditions. It combines the functions of physical training and rehabilitation training. The acceleration signal changes to a certain extent to reflect the changes of the human dumbbell exercise, so the main frequency range of the acceleration signal is mainly distributed in the range of 0–10 Hz. FFT transformation is performed on the collected acceleration signal and its spectrum is viewed, as shown in Figure 4. It can be found that the main frequency components are within 5 Hz, and the amplitude of the signal above 5 Hz is almost negligible. The acceleration signal is low-pass filtered to retain the signal components from 0 to 10 Hz and remove the effect of high-frequency noise.

In this study, second-order low pass filtering will be used to filter the signal. The expression of the second-order filter transfer function is given by the following equation:

$$d \cdot \frac{d^2 y(t)}{dt^2} - e \cdot \frac{dy(t)}{dt} - c_2 y(t) = a \cdot \frac{d^2 x(t)}{dt^2} - b \cdot \frac{dx(t)}{dt} - c_1 x(t). \tag{8}$$

$y(t)$ is the output signal, and $x(t)$ is the input signal. For a dumbbell movement that rotates around an axis, such as a bending movement, the curve with the largest variance among the three-axis angular velocity curves
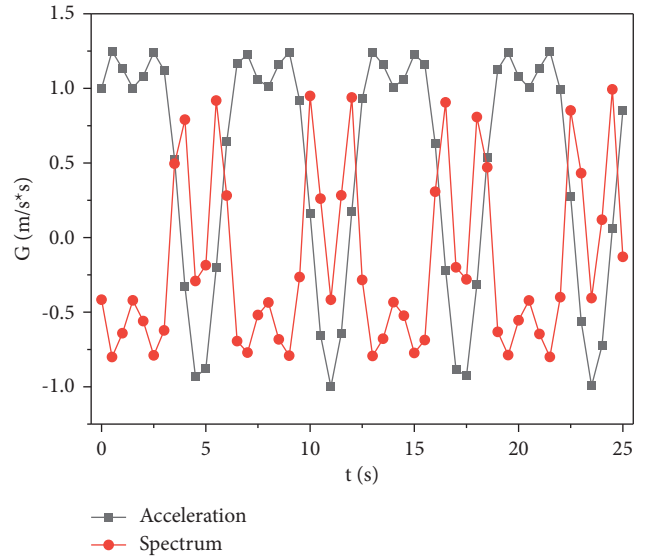
FIGURE 4: Acceleration signal spectrum.

collected at the acquisition end is selected as the judgment curve of the movement endpoint. The over-zero point of the curve is the highest or lowest point of a single dumbbell action, and the curve between the two zero points is a cycle curve of the dumbbell action. Since the collected data are discrete points and not a continuous curve, and these discrete data cannot be well fitted using a sine function, it is not possible to directly find out the over-zero point of the angle curve. In this study, a moving average filter and an energy threshold judgment are introduced to find the valid angle curve over zero points. The period of a single dumbbell action is about 3 seconds and the sampling frequency is 100 Hz, we can know that there are about 300 discrete points in the angle curve for each dumbbell action.

2.1.3. Experimental Design for Physical Training Moderation.

Evaluation is the key link of the digital control process, the setting, and modification, and improvement of training objectives, the formulation, implementation, and regulation of training plan, the evaluation, identification, and feedback of training results all need the participation of evaluation system. Making full use of the guidance, identification, feedback, and motivation of evaluation indexes, a complete set of special physical quality evaluation systems for javelin throwers is established to evaluate the physical fitness level of athletes, such as avoiding the one-sidedness caused by the evaluation of individual physical quality indexes and comparing the gap between the strong and weak items of individual physical quality [19]. According to the average price result, we can carry out targeted training to improve the athletes' physical fitness level. Physical fitness evaluation is "the assessment of the real state of the athlete's physical movement ability."

The comprehensive level of physical fitness in this study refers to the average level of 10 physical fitness indicators of athletes. As the athletes' fitness level improves, the growth rate of each physical quality gradually decreases and the athletes' pursuit of physical quality gradually fades. Long-term accumulation may lead to athletes' shortcomings in physical fitness and gradually widen the distance between them and excellent athletes, which is not conducive to athletes' long-term development. The sample mutation is guided by the timely use of the algorithm to achieve the purpose of the present fuzzing test effect. In response to the above situation, the comprehensive evaluation standard of physical fitness applied to evaluate athletes comprehensively, so that the overall physical fitness level of athletes can be steadily improved and the concept of reflecting the overall level by single quality is rejected, as listed in Table 2.

The specificity of basketball sports is the need for athletes to have the strong physical strength to accomplish the embodiment of technical movements. In the study, no significant variability was found in the relative peak rotational torque of hip and knee flexors and extensors on both sides in our male basketball players. The results showed that knee strength and knee joint strength were more balanced, and there was no significant difference between the left and right-side surfaces [20].

The second phase is the competition period. It can be observed that in the second phase, with the intervention of this training method, the player's basic strength does not decline, but still maintains a small growth, but not as much as in the first phase, which can be observed by the T-value less than -10. The goal of Phase II is to maintain the player's base strength capacity so that it does not decrease due to the high level of competition and exertion during the competition, allowing the player to maintain peak physical condition for the entire season.

3. Results and Analysis

3.1. Performance Results of Wavelet Recursive Fuzzy Neural Network Algorithm.

The two-axis position synchronization error when using a global sliding mode control method

Table 2: Paired sample check.

Pair 1	Step 1: weight pre
	Step 2: weight post
Pair 2	Step 1: one-step vertical take-off pre
	Step 2: one-step vertical take-off post
Pair 3	Step 1: vertical take-off pre
	Step 2: vertical take-off post
Pair 4	Step 1: bench press 1RM pre
	Step 2: bench press 1RMpost squat 1RM pre
Pair 5	Step 1: squat 1RM postmaximum number of 60 kg bench press
Pair 6	Step 1: maximum number of 60 kg bench press post
Pair 7	Step 1: maximum number of pull-ups post
	Step 2: maximum number of pull-ups post

combined with a recursive wavelet fuzzy neural network synchronization compensator. Compared with the cross-coupled controller, the maximum value of the dual-axis position synchronization error is reduced from $16.78\,\mu m$ to $1.02\,\mu m$, and the maximum magnitude during regulation is $15.63\,\mu m$. Using the recursive wavelet fuzzy neural network synchronization compensator designed in this chapter between the two axes can effectively improve the anti-interference performance of the platform compared to the cross-coupled control compensator.

To facilitate the extraction of feature values, 50,000 sets of seems signals collected from the plate support exercise were analysed by grouping every 200 data points into one group, with a total of 250 sets of data, using the first 80% sets of data for network training and the last 20% sets of data for prediction. The prediction results are shown below, where the x-axis represents the data groups and the y-axis represents the frequency values of MF (MPF). The data fitting curve is shown in Figure 5.

The recursive wavelet fuzzy neural network synchronization compensator designed in this study improves the synchronization performance of the dual axes in the Y direction compared with the cross-coupling control compensator, and it can instantly adjust the compensation signal of the dual axes under the change of the beam load, improve the tracking accuracy of the system in the single axis, suppress the influence of the sudden added load on the system, and make the system have a better antidisturbance performance, including plan formulation, training data statistics, stage training effect evaluation, and stage training effect feedback.

If adaptive adjustment of the size of the parameters ensures that the sum of squares of the network errors is minimized, network training can greatly improve the generalization ability of the network by effectively controlling the complexity of the network. During the neural network training, a correction function for the performance function is introduced based on the conventional mean square error performance function. The performance of the neural network is analysed by comparing the mean square error MSE before and after optimization, and the results are shown in Figure 6.

The network performance indexes obtained by LM and SCG optimization methods are better than the original
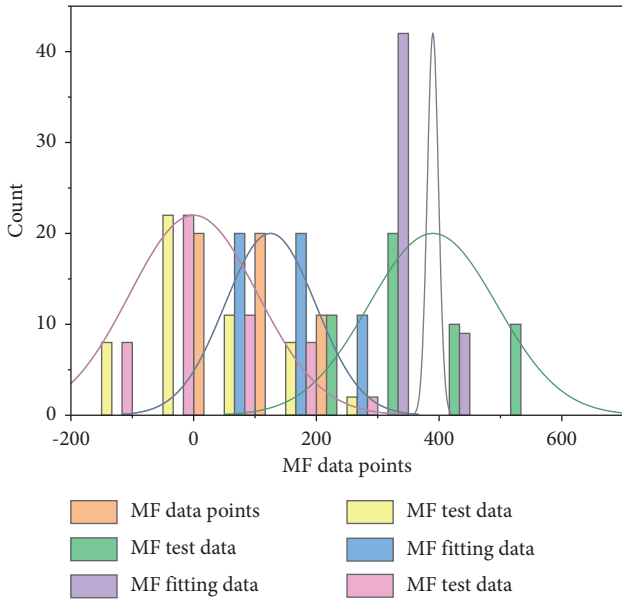
FIGURE 5: Neural network prediction curve fit for MPF values.

FIGURE 6: Mean square error MSE of MF before and after optimization for each channel.

indexes, in which the LM algorithm is the best, it is known from the quantitative data that the application of the LM optimization algorithm has a great improvement in the accuracy of the designed BP network, and it is also more accurate for the prediction effect, which reduces the prediction error and achieves the purpose of optimization.

FIGURE 7: Experimental results.

The simulation results show that the recursive wavelet fuzzy neural network compensation controller can dynamically compensate the single-axis tracking error and the inter-axis synchronization error of the system so that the tracking error and synchronization error of the system are effectively reduced and the influence of the load change and uncertain disturbance terms on the control process of the system is reduced, thus improving the control accuracy and robustness of the H-type motion platform.

3.2. Experimental Results of Real-Time Modulation of Physical Training Intensity. When designing a training program, the formulation of phased training objectives and the evaluation of training results are of great importance. The development of training objectives is the premise and basis for the entire training process, and the evaluation of training results is a key step in testing the training program. The combination of these two can be used to test the effectiveness of the current phase of training and to further grasp the overall situation of each athlete's training effectiveness throughout the year.

Compare the "actual training results" with the "expected target results," if there is no significant difference between the two, then the original training plan is continued to implement; if the difference between the two is more obvious, then the reasons for the deviation are analysed, timely adjustment of the training plan is carried out, and different feedback means are taken. In the following section, we will analyse the completion of the target in the preparation period, the first competition period, and the second competition period, as well as the reasons for the deviation, and point out the adjustment measures to "correct the deviation," as shown in Figure 7.

If all the training data of the athlete is normal, and the training intensity and training volume meet the training

requirements, but the deviation still exists and is outside the error allowance, the reason is most likely that the goal is set too high. The second is that the athlete fails to maintain a good view of training. This is manifested by the abnormal training data of a certain athlete, where the load intensity, load volume, and target completion rate are significantly lower than those of other athletes. Negative ones indicate reverse inhibition. The connection weight is like the ganglion in the actual neuron structure, which is used to indicate the connection strength between two nerve cells. After continuous learning and adjustment, the output of the neuron can gradually approach the required expected value.

Coaches should look at the athletes' strengths and weaknesses from a "developmental" perspective and combine "multi-year training goals," "annual training goals," "large cycle training goals," and so on. The coach should design a long-term training plan, continue to develop the athlete's physical strengths, and gradually adjust the athlete's shortcomings. Physical training needs to be a combination of work and rest, and a good balance between training and recovery time.

4. Conclusion

In the past, sports medicine has focused on physiologically based recovery and rehabilitation but has neglected to focus on rehabilitation physical training and injury prevention physical training. This has led to repeated injuries and long recovery times for basketball players, as well as an increased risk of injury unknowingly because of injury and lack of indepth knowledge of physical training. There are major shortcomings in fitness testing, program development, training arrangements, training methods, training monitoring, and effect evaluation, which are the main reasons for the high incidence of sports injuries. Through the assessment and diagnosis of athletes' physical fitness status and real-time monitoring, the training quality and training load control of physical training can be ensured to effectively reduce the risk of sports injuries. In addition, through the eight basic injury prevention tests, athletes can be targeted to improve their ability to prevent sports injuries and to maintain their athletic performance through functional training and sports medicine treatment programs after suffering a sports injury, and to recover their original athletic performance in a short period and then to improve it. Extracting and analysing the eigenvalues in the time and frequency domains, a multilayer algorithm incorporating decision trees were designed to successfully identify 8 movements. The experimental results show that the algorithm achieves a combined recognition rate of 96.3%.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] A. Ashrafian, M. Mirsalim, and M. A. Masoum, "An adaptive recursive wavelet based algorithm for real-time measurement of power system variables during off-nominal frequency conditions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 818–828, 2017.

[2] S. Ranjan Das, A. K. Mishra, P. K. Ray et al., "Advanced wavelet transform based shunt hybrid active filter in PV integrated power distribution system for power quality enhancement," *IET Energy Systems Integration*, vol. 2, no. 4, pp. 331–343, 2020.

[3] K. R. Reddy, "Power quality classification of disturbances using discrete wavelet packet transform (DWPT) with adaptive neuro-fuzzy system," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 3, pp. 4892–4903, 2021.

[4] S. Ganjefar, M. Afshar, M. H. Sarajchi, and Z. Shao, "Controller design based on wavelet neural adaptive proportional plus conventional integral-derivative for bilateral teleoperation systems with time-varying parameters," *International Journal of Control, Automation and Systems*, vol. 16, no. 5, pp. 2405–2420, 2018.

[5] Y. Zhu, Q. Yan, and J. Lu, "Fault diagnosis method for disc slitting machine based on wavelet packet transform and support vector machine," *International Journal of Computer Integrated Manufacturing*, vol. 33, no. 10-11, pp. 1118–1128, 2020.

[6] S. Nanda and P. K. Dash, "Field programmable gate array implementation of fuzzy variable step size adaptive linear element for adaptive frequency estimation," *IET Signal Processing*, vol. 11, no. 9, pp. 1083–1094, 2017.

[7] Y. Chen, Q. Cheng, Y. Cheng, H. Yang, and H. Yu, "Applications of recurrent neural networks in environmental factor forecasting: a review," *Neural Computation*, vol. 30, no. 11, pp. 2855–2881, 2018.

[8] P. Kumar and A. S. Hati, "Review on machine learning algorithm based fault detection in induction motors," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1929–1940, 2021.

[9] D. C. dos Santos Gomes and G. L. de Oliveira Serra, "Computational approach for real-time interval type-2 fuzzy kalman filtering and forecasting via unobservable spectral components of experimental data," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 2, pp. 326–355, 2021.

[10] X. He, J. Luo, G. Zuo, and J. Xie, "Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks," *Water Resources Management*, vol. 33, no. 4, pp. 1571–1590, 2019.

[11] L. Han, R. Zhang, X. Wang, A. Bao, and H. Jing, "Multi-step wind power forecast based on VMD-LSTM," *IET Renewable Power Generation*, vol. 13, no. 10, pp. 1690–1700, 2019.

[12] H. Liu and L. Shen, "Forecasting carbon price using empirical wavelet transform and gated recurrent unit neural network," *Carbon Management*, vol. 11, no. 1, pp. 25–37, 2020.

[13] C. Shen, Y. Zhang, X. Guo et al., "Seamless GPS/inertial navigation system based on self-learning square-root cubature Kalman filter," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 1, pp. 499–508, 2020.

[14] R.-E. Precup, T.-A. Teban, A. Albu, A.-B. Borlea, I. A. Zamfirache, and E. M. Petriu, "Evolving fuzzy models for prosthetic hand myoelectric-based control," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4625–4636, 2020.

[15] M. Li, W. Chen, and T. Zhang, "FuzzyEn-based features in FrFT-WPT domain for epileptic seizure detection," *Neural Computing & Applications*, vol. 31, no. 12, pp. 9335–9348, 2019.

[16] I. S. Samanta, P. K. Rout, and S. Mishra, "An optimal extreme learning-based classification method for power quality events using fractional Fourier transform," *Neural Computing & Applications*, vol. 33, no. 10, pp. 4979–4995, 2021.

[17] O. Avalos, E. Cuevas, H. G. Becerra, J. Gálvez, S. Hinojosa, and D. Zaldívar, "Kernel recursive least square approach for power system harmonic estimation," *Electric Power Components and Systems*, vol. 48, no. 16-17, pp. 1708–1721, 2021.

[18] M. Zhang, X. Liu, Z. Zhang et al., "A novel modeling approach and its application in polymer quality index prediction," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 7, pp. 2005–2015, 2019.

[19] F.-J. Lin, S.-G. Chen, S. Li, H.-T. Chou, and J.-R. Lin, "Online autotuning technique for IPMSM servo drive by intelligent identification of moment of inertia," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7579–7590, 2020.

[20] A. K. Adik and W. Wang, "An intelligent system for real-time condition monitoring of tower cranes," *Intelligent Control and Automation*, vol. 10, no. 4, pp. 155–167, 2019.

Research Article

Analysis of Spatial Distribution of CVD and Multiple Environmental Factors in Urban Residents

Beichen Wang,[1] Kangkang Gu [1], Dong Dong,[1] Yunhao Fang,[2] and Lingling Tang[3]

[1]*School of Architecture & Planning, Anhui Jianzhu University, Hefei, Anhui 230022, China*
[2]*School of Architecture and Urban Planning, Huazhong University of Science and Technology, Wuhan, Hubei 430000, China*
[3]*College of Architecture and Art, Hefei University of Technology, Hefei, Anhui 230009, China*

Correspondence should be addressed to Kangkang Gu; kangkanggu@ahjzu.edu.cn

Cardiovascular disease (CVD) poses a serious threat to urban health with the development of urbanization. There are multifaceted and comprehensive influencing factors for CVD, so clarifying the spatial distribution characteristics of CVD and multiple environmental influencing factors is conducive to improving the active health intervention of urban environment and promoting the sustainable development of cities The spatial distribution characteristics of CVD deaths in a certain district, Bengbu City, Huaihe River Basin, China, in 2019 were explored, and the correlation between multiple environmental factors and CVD mortality was investigated in this study, to reveal the action mechanism of multiple environmental factors affecting the risk of mortality. Relevant studies have shown that (1) CVD deaths are characterized as follows: male deaths are more than females; the mortality is higher in those of higher age; most of them are unemployed; cardiocerebral infarction is the main cause of death; and the deaths are mainly distributed in the central city and near the old urban area. (2) The increased CVD mortality can be attributed to the increased density of restaurants and cigarette and wine shops around the residential area, the increased traffic volume, the dense residential and spatial forms, the low green space coverage, and the distance from rivers. Therefore, appropriate urban planning and policies can improve the active health interventions in cities and reduce CVD mortality.

1. Introduction

A great number of studies have demonstrated the high mortality and disability rate of cardiovascular disease (CVD) [1] and the resulting global burden of diseases [2]. The traditional risk factors for CVD include hypertension, diabetes, dyslipidemia, obesity, smoking, alcohol consumption, and lack of physical exercise. However, under the biological-psychosocial medical model, the influencing factors on health are multifaceted and comprehensive, including environmental factors. Previous studies have shown that the urban population is more likely to suffer from cardiovascular diseases than the rural population in some developing countries [3]. Industrialization, urbanization, aging, and constant changes in the ecological environment and lifestyle have brought challenges to people's health in the rapid urban development. A lot of cities are experiencing a major shift in the disease spectrum from infectious diseases to chronic noncommunicable diseases. Therefore, it is conducive to improving active health interventions in urban environment and promoting sustainable urban development in the future to identify the mortality distribution characteristics of CVD and environmental influencing factors.

Previous studies have shown that unreasonable diet, excessive oil, and salt intake will increase the pressure of human circulatory system [4], thereby leading to CVD morbidity and mortality. The increase in the amount of physical activity, such as work, commuting, and leisure, will appropriately reduce the risk of CVD-related death [5], and the accessibility of sports and leisure activity facilities and open space will also have an influence on the level of physical activity of residents [6]. Residents will tend to commute by non-motor vehicles and increase daily activities provided that the urban road slow traffic system is perfect and the bus

stops are dense [7]. In addition, residents engaged in manual work will have increased daily activities [8], thus reducing the risk of death. Air pollution [9–11] and noise pollution [12] constitute the important risk factors for CVD, of which long-term air pollution has a greater impact on CVD than short-term one [13]. The closer distance to the pollution source or the greater density of the surrounding pollution sources will increase the possibility of exposure to pollutants [14, 15], including a large number of industrial pollutants [16], the cooking fumes from [17], the exhaust gas emissions of motor vehicles, and urban transportation facilities, such as railways, subway lines, and airports [18]. Roads and restaurants produce noise pollution while causing air pollution, thereby leading to the occurrence of cardiovascular diseases [19]. The nearer distance from rivers and the higher green space coverage can effectively reduce the concentration of particulate matter (Douglas, W., and [20–22]) and lower CVD mortality. Cold weather and sudden temperature drop in winter are prone to causing overreaction of human function, resulting in death from acute CVD [23, 24]. The difference in temperature in different urban regions will lead to the difference in mortality.

In general, current studies focus on the macro-level of county [25], prefecture [26], and state [25], but neglect the small scales, such as communities and streets; meanwhile, the consideration of environmental factors is relatively simple, and the influence of aging population and old urban areas is relatively ignored. A large number of people in the Huaihe River Basin in Central China died of chronic diseases caused by the environment many years ago [27]. The urban environmental pollution is relatively serious now, seriously endangering the physical and mental health of residents. In view of this, the distribution characteristics of CVD deaths were analyzed, and the influences of multiple environmental factors on CVD were systematically investigated in this study with the old urban area in Bengbu City of Huaihe River Basin as the subject, to reduce CVD mortality by proper urban planning and policies, and provide certain research ideas for healthy city construction in other areas of Huaihe River Basin.

2. Materials and Methods

2.1. Study Area.
Bengbu City, located in the south of Huaihe River Basin, China, is characterized by high population density and rapid urbanization [28]. Bengbu was an important industrial city and traffic hub in China before, where a large number of chemical, pharmaceutical, and building materials enterprises and the heavy traffic pressure have produced a large amount of pollution in Bengbu, seriously affecting the physical and mental health of residents. A certain district is located in the Western old urban area of Bengbu City, adjacent to the commercial and cultural center of the city in the east and the Huaihe River (one of the seven major rivers in China) in the north, and it is also the administrative area with the largest population in Bengbu City, so it is suitable for the research on urban environment and population health. The study area consisted of 12 blocks numbered YH01-YH12 in a certain district, which were divided into 2,744 spatial units by 100 m grid (Figure 1). The study area is located in the mid-latitude region, where northeasterly wind is the most prevalent throughout the year, the average temperature in January is 1°C, and it is susceptible to cold waves in winter.

2.2. Data Sources

2.2.1. Sources of Data Related to CVD.
The data of cardiovascular disease (CVD) mortality were collected from the population whose cause of death was cardiovascular disease (ICD-10 code I00-I99) determined by Bengbu Health Commission in 2019, including detailed address, age, gender, and occupation before death. Inpatients whose home addresses and affiliations were not within the study area were excluded, so the final number was 469 in total. The number of resident population in a certain district in 2019 was derived from the official bulletin of Bengbu National Bureau of Statistics.

2.2.2. Environmental Data Sources.
The land use data and built environment data in environmental data were obtained from the Anhui Urban & Rural Planning and Design Institute, and the data contained the main uses of each land in a certain district in 2019, as shown in Figure 2(a), which included roads, green space and river, residential land, industrial land according to urban land classification, and the base area and floor number of each building. Commercial and bus stops are derived from open-source point-of-interest (POI) data, which are characterized by large sample size, wide coverage, and detailed spatial resolution, making spatial analysis more comprehensive, objective, and in-depth, as shown in Figure 2(b). Data of environmental pollution indexes were obtained from the Bengbu Bureau of Ecology and Environment, and the pollution values of each spatial unit were obtained by the interpolation method for the pollution monitoring point data. Data of land surface temperature and vegetation coverage were derived from Landsat 8 remote sensing data in January 2019 provided by USGS and coded LC81210372019023LGN00.

2.3. Influencing Factor Index System.
To explore the distribution characteristics of cardiovascular disease (CVD) and environmental influencing factors in urban residents, the basic characteristics and spatial distribution characteristics of CVD deaths were analyzed, and the correlation between various environmental influencing factors and CVD mortality was explored (Figure 3). Environmental influencing factors included land use, road traffic, spatial form, and natural environment, with their potential impacts on CVD shown in Table 1. The calculation formula for CVD mortality is as follows:

$$\text{Mortality} = \frac{\text{Nc}}{\text{Np}} \times 100\%, \tag{1}$$

where Nc is the number of deaths due to CVD in a unit and Np is the total population in a unit.
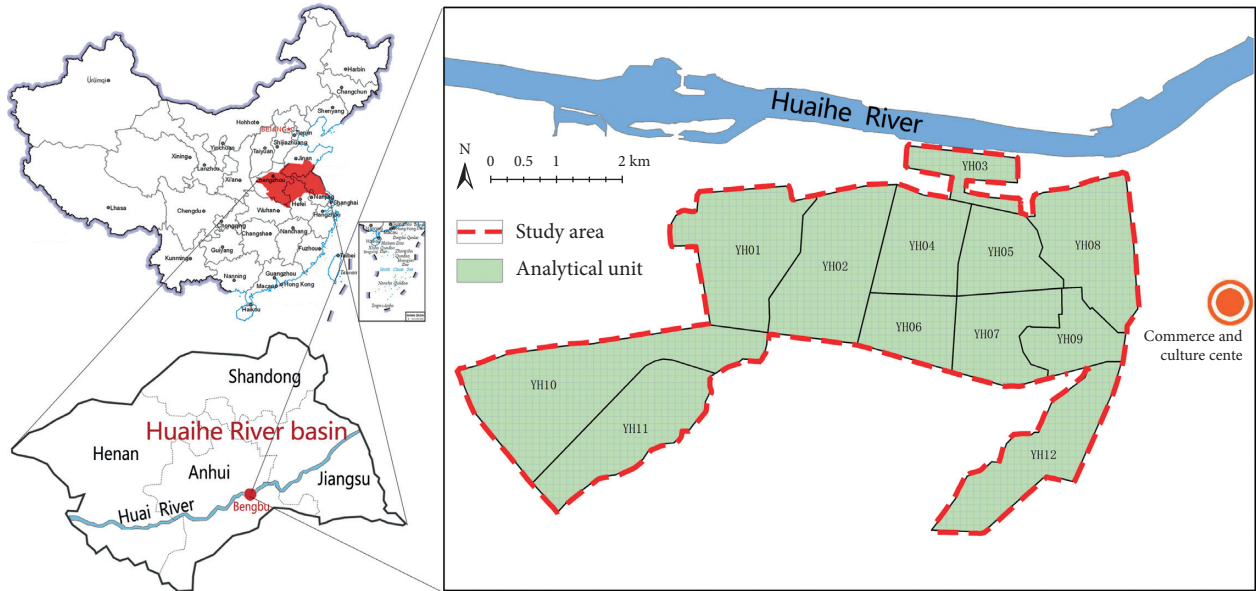
Figure 1: Area bitmap.

(a)   (b)
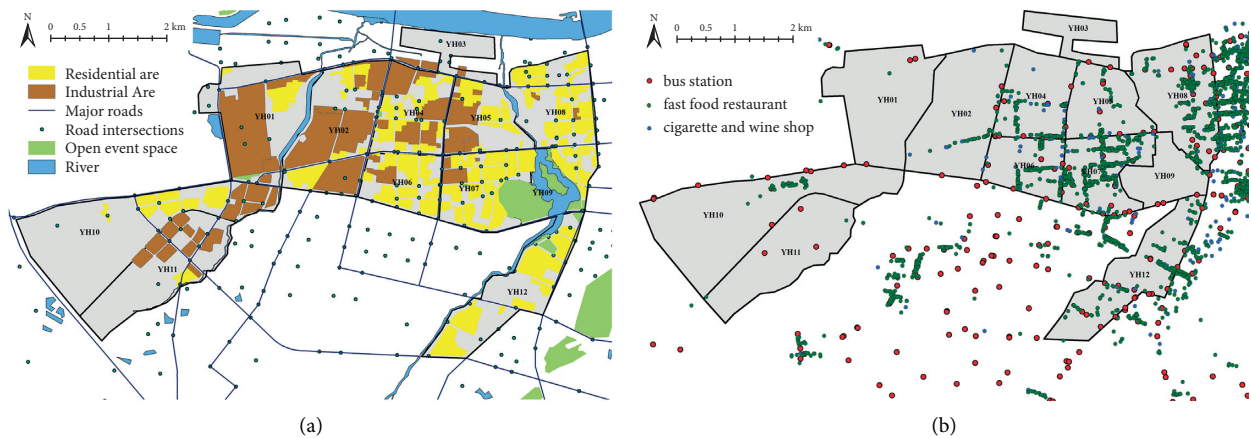
Figure 2: Present distribution.
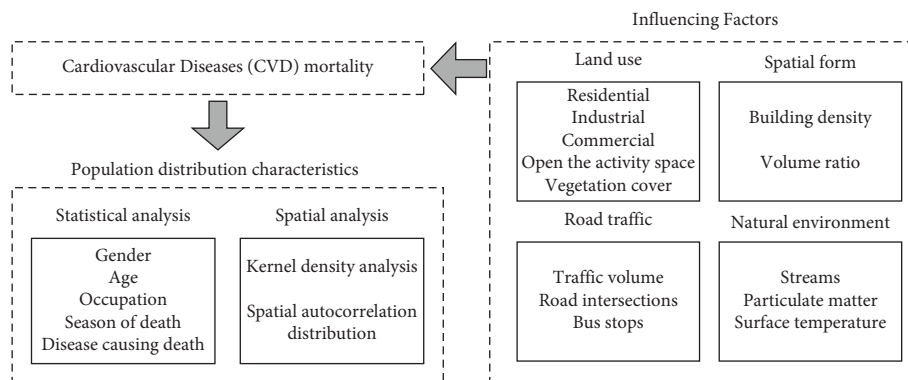
Figure 3: Conceptual framework.

2.4. Study Methods

2.4.1. Spatial Autocorrelation Assessment. Global spatial autocorrelation [29] is usually measured by Moran's I, and

the aggregation of subregional geospatial CVD deaths is judged according to the spatial distribution pattern of the whole region. Moran's I range is $[-1, 1]$, and the larger value indicates the more obvious spatial correlation, and the

TABLE 1: Environmental influencing factors.

Category	Subcategory	Variable	Unit	Definition	Potential impact of hypothesis
Land use	Live	Residential density	%	Living area/total land area	Cause pollution and physical and mental stress
	Industry	Distance from the nearest industry	m	Distance from the nearest industry	The closer you are to industry, the more pollution you get
	Business	Fast-food restaurant density	Pcs/ha	Number of fast-food restaurants in 500 m buffer zone/buffer zone area	Intake relatively high salt and oil
		Tobacco and liquor shop density	Pcs/ha	Number of tobacco and liquor stores in 500 m buffer zone/buffer zone area	Increase the risk of alcohol and smoking
	Open the activity space	Distance from the nearest open space	m	Distance from the nearest open space	It is more convenient to exercise and relax
	Vegetation cover	Normalized difference vegetation index (NDVI)	—	NDVI in 2019	Vegetation coverage can reduce air pollution, increase green vision rate, and delight body and mind
Road transport	Traffic	Road network density	m/ha	Road length/total land area	Increase air and noise pollution
	Road intersection	Road intersection density	Pcs/ha	Number of road intersections in 500 m buffer zone/buffer zone area	More air and noise pollution
	Bus stop	Bus stop density	Pcs/ha	Number of bus stops in 500 m buffer zone/buffer zone area	It is more convenient to travel and increase the amount of activities
Spatial form	Building density	Building coverage	%	Total area of building basement/total area of land	High density reduces the wind speed, which increases the concentration of air pollutants and makes people depressed
	Volume fraction	Volume fraction	—	Total construction area/total land area	High density reduces the wind speed, which increases the concentration of air pollutants and makes people depressed
Natural environment	River	Distance from the nearest river	m	Distance from the nearest river	Adsorption of air particles by water body
	Particulate matter (PM) (Ground) surface temperature	Mean PM$_{2.5}$	Ug/m^3	Average PM$_{2.5}$ in 2019	Increase air pollution
		Mean PM$_{10}$	Ug/m^3	Average PM$_{10}$ in 2019	Increase air pollution
		Mean surface temperature	°C	Surface temperature in January 2019	Cold winter temperatures increase the risk of death

smaller value indicates the greater spatial difference, with the specific calculation formula as follows:

$$\text{Moran's } I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n} w_{ij} \sum_{j=1}^{n} \sum_{i=1}^{n} (x_i - \overline{x})^2} \ (i \neq j), \qquad (2)$$

where n is the number of study units; x_i and x_j represent the CVD mortality in area i and area j, respectively. $\overline{x}$ represents the average CVD mortality; W_{ij} represents the spatial weight matrix.

Local spatial autocorrelation [30] can reveal the heterogeneity of local space, make up for the atypical characteristics of local areas in the whole region, and fully reflect the variation trend of local geospatial epidemic risk. The spatial difference in CVD mortality was measured by local indicators of spatial association (LISA) in this study, with the calculation formula shown as follows:

$$\text{LISA}_i = z_i \sum_j w_{ij} z_j, \qquad (3)$$

where z_i and z_j represent the standardized values of CVD mortality in area i and area j, respectively. W_{ij} represents the

spatial weight matrix. In the LISA of local spatial autocorrelation, CVD mortality can be divided into four categories: high-high clustering (H-H), high-value clustering; low-low clustering (L-L), low-value clustering; high-low clustering (H-L), outliers with high values mainly surrounded by low values; and low-high clustering (L-H), outliers with low values mainly surrounded by high values.

2.4.2. Spearman's Rank Correlation. To investigate the relationship between CVD and environmental factors, the correlation between the mortality rate of CVD (dependent variable) and the values of various environmental factors (independent variables) was comprehensively measured by Spearman's rank correlation [31], with the specific formula referred to formula (3). Compared with Pearson's correlation, Spearman's rank correlation can make up for the fact that the correlation of environmental factor values may not be linear and are interfered with detection error and other factors, thus making the correlation more significant [32]. CVD mortality and environmental factor values were subject

to normalized processing (value range of [0, 1]), which was used to balance the dimensional gap of data and enable different data to be counted under the same conditions.

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}, \quad (4)$$

where ρ is Spearman's correlation coefficient, with the value range of [−1, 1], and the larger absolute value indicates the stronger correlation. n represents the number of areas of CVD mortality, and d_i represents the rank difference between dependent variable (CVD mortality) and independent variable (environmental influencing factor values).

3. Results

3.1. Basic Characteristics of CVD Deaths. A total of 469 people died of cardiovascular disease (CVD) in the study area in 2019 according to the statistics of collected death data, with the basic characteristics of death population, such as gender, age, occupation, death season, and specific diseases leading to death, shown in Table 2.

The analysis results showed that the mortality among male CVD patients was significantly higher than female patients, with a ratio of about 1.18 : 1. The elderly over 80 years accounted for 54.15% of CVD deaths, and those under 60 years only accounted for about 10%, indicating that the higher the age, the higher the mortality. Winter (December to February of the following year) and autumn (September to November) were the most frequent periods of CVD deaths, accounting for 27.93% and 27.72% of the total number of deaths, respectively, and the mortality in autumn and winter was 1.25 times higher than that in spring and summer. The largest number of CVD deaths was among the unemployed (26.01%), followed by workers (6.61%), freelancers (5.54%), and farmers (2.56%). Among the specific diseases causing death, cerebral infarction and myocardial infarction accounted for 24.09% and 19.62%, followed by coronary heart disease (17.70%).

3.2. Spatial Distribution of CVD Deaths. In 2019, there was a significant difference in the number of CVD deaths among different units in the study area. YH07 and YH08 accounted for 73.56% in terms of deaths, of which YH08 had the largest number of CVD deaths, namely 180, followed by 165 deaths in YH07 and no deaths in YH10 and YH11. The kernel density analysis of CVD deaths was conducted with ArcGIS software (Figure 4), and unit density of CVD deaths in the study area could be calculated, which could intuitively reflect the distribution of CVD deaths in a continuous area [33]. The analysis results suggested that CVD deaths were mainly distributed in the vicinity of Unit YH09 in Zhanggongshan Park and the north of Unit YH08 near the Huaihe River. The area near YH09 was an old urban area of Bengbu, characterized by high population density and serious aging. A total of 365 people died within the 1 km buffer zone of

Zhanggongshan Park, accounting for 77.82% of the total deaths; moreover, the density of deaths was lower in the area farther away from Zhanggongshan Park. The north of YH08 is also a large population gathering area, close to the central city of Bengbu and Huaihe River in the east.

Among 2,744 study units, 1,053 units containing residential land (i.e., unit population greater than 0) were taken as the research samples after the outlier samples were excluded. The normalized mean value of the calculated CVD mortality values of all study units was normalized to be 0.008, with the standard deviation of 0.055, indicating that there were significant spatial differences in the CVD mortality. Global Moran's I index analysis showed that the spatial distribution of normalized mortality index was of spatial autocorrelation (Moran's $I > 0$, $Z = 2.192$, $P < 0.05$), indicating that there was a significant spatial aggregation phenomenon in the distribution of CVD deaths. In addition, there was a local spatial autocorrelation in the normalized mortality index according to the spatial distribution of LISA clustering (Figure 5), in which H-H clustering (22) was mainly distributed in the middle of YH08, forming a significant H-H clustering of CVD mortality. L-L clusters (121) were mainly distributed in YH05, YH10, and YH12, with a few scattered clusters in other units, showing more L-L clusters in the CVD mortality.

3.3. Relationship between Environmental Factors and CVD Mortality. The results of Spearman's rank correlation analysis are shown in Table 3. Among land use, road traffic, spatial form, and natural environment, there are 14 environmental factors showing significant correlation and statistical significance with CVD mortality ($P < 0.05$). There was no correlation between industry and CVD mortality ($P > 0.05$). The spatial distribution of environmental influencing factors is shown in Figure 6.

3.3.1. Land Use. The model showed that the residential density was positively correlated with the CVD mortality ($\rho = 0.127$), indicating that the higher the residential density, the higher the CVD mortality. The density of fast-food restaurants and the density of cigarette and wine shops showed a significant positive correlation with the CVD mortality in the 500 m buffer zone ($\rho = 0.287$, $\rho = 0.286$), indicating that the higher the density of fast-food restaurants and cigarette and wine shops, the higher the mortality. There was a significant negative correlation between the nearest distance to the open space and the CVD mortality ($\rho = -0.189$), indicating that the farther from the park, the lower the CVD mortality. The normalized difference vegetation index (NDVI) in January showed a significant negative correlation with CVD mortality ($\rho = -0.179$), indicating that the higher the vegetation coverage, the lower the CVD mortality.

3.3.2. Road Traffic. The model showed that the road network density and the intersections within the 500 m buffer zone were significantly positively correlated with the CVD

TABLE 2: Statistics of CVD death population.

Category	Classify	Death toll	Proportion of deaths (%)
Gender	Man	254	54.16
	Woman	215	45.84
Age	≥80 years old	254	54.15
	70–79 years old	109	23.24
	60–69 years	58	12.37
	<60 years old	48	10.23
Occupation	Unemployed	122	26.01
	Worker	31	6.61
	Individual operator	26	5.54
	Farmer	12	2.56
	Other	278	59.27
Dead season	Spring (March–May)	110	23.45
	Summer (June–August)	98	20.90
	Autumn (September–November)	130	27.72
	Winter (December–February of the following year)	131	27.93
Specific diseases leading to death	Cerebral infarction	113	24.09
	Cardiac infarction	92	19.62
	Coronary heart disease	83	17.70
	Encephalorrhagia	55	11.73
	Pulmonary heart disease	43	9.17
	Other	83	17.70
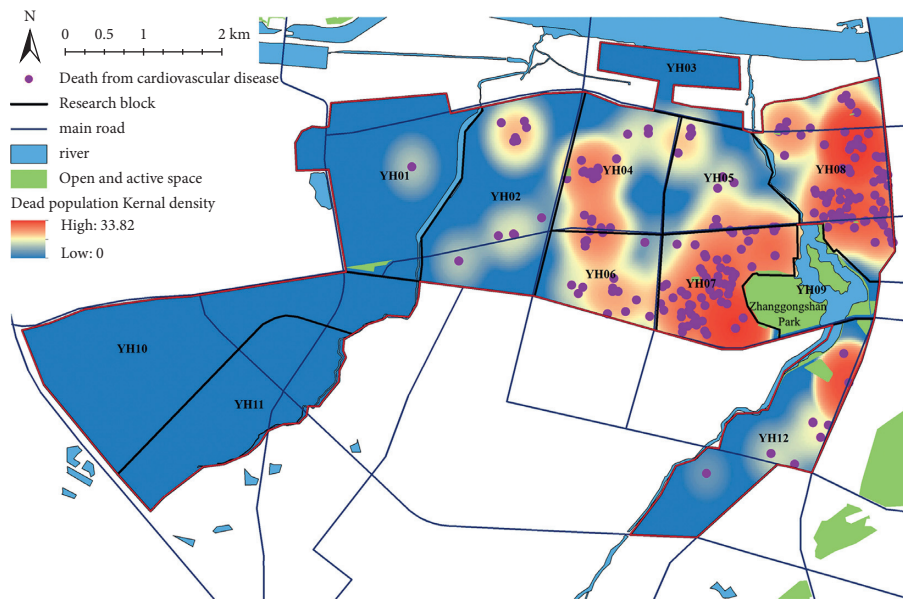
FIGURE 4: Distribution map of CVD death population.

mortality ($\rho = 0.138$, $\rho = 0.220$), indicating that the greater the traffic volume, the higher the CVD mortality. Meanwhile, there was a significant positive correlation between the density of bus stops and the CVD mortality within the 500 m buffer zone ($\rho = 0.238$), indicating that there was a lower risk of death among the people living in the area with sparse bus stops.

3.3.3. Spatial Form. The model showed that building density and floor area ratio were significantly positively correlated with CVD mortality, with the correlation coefficients of 0.104 and 0.115, indicating that high-density space would increase the CVD mortality.

3.3.4. Natural Environment. There was a significant positive correlation between the distance to the nearest river and the CVD mortality ($\rho = 0.121$), indicating that the closer to the river, the lower the CVD mortality. The average annual $PM_{2.5}$ and PM_{10} were significantly positively correlated with the CVD mortality, with the correlation coefficients of 0.088 and 0.072, indicating that the higher the average annual particulate matter (PM) concentration, the higher the CVD
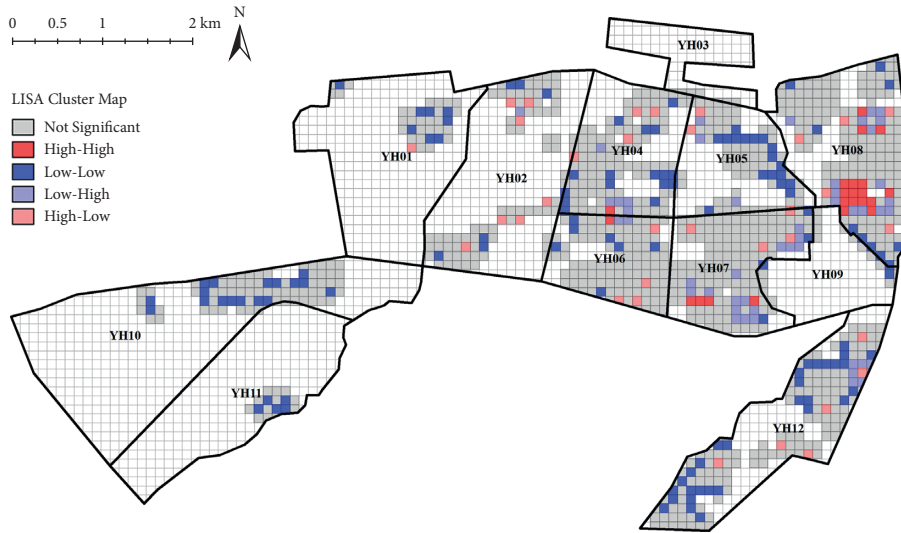
FIGURE 5: Spatial autocorrelation distribution of CVD normalized mortality index.

TABLE 3: Correlation between environmental factors and CVD mortality.

Category	Subcategory	Variable	P (correlation coefficient)	P value (double side)
	Live	Residential density	0.127*	0.000
	Business	Density of fast-food restaurants in 500 m buffer zone	0.287**	0.000
		Density of tobacco and liquor stores in 500 m buffer zone	0.286**	0.000
Land use	Industry	Distance from the nearest industry	-0.011	0.717
	Open the activity space	Distance from the nearest open space	−0.189**	0.000
	Vegetation cover	Normalized difference vegetation index (NDVI)	−0.179**	0.000
Road transport	Traffic	Road network density	0.138**	0.000
		Road intersection density in 500 m buffer zone	0.220**	0.000
	Bus stop	Density of bus stops in 500 m buffer zone	0.238**	0.000
Spatial form	Building density	Building coverage	0.104**	0.001
	Volume fraction	Volume fraction	0.115**	0.000
	River	Distance from the nearest river	0.121**	0.000
Natural environment	Particulate matter (pm)	Average annual $pm_{2.5}$	0.088**	0.004
		Average annual pm_{10}	0.072*	0.020
	(Ground) surface temperature	Surface temperature in January 2019	−0.135*	0.000

* $P < 0.05$, ** $P < 0.01$.

mortality. The average temperature was negatively correlated with the CVD mortality, with a correlation coefficient of -0.135, indicating that the lower the temperature in winter would contribute to the increase in the CVD mortality.

4. Discussion

(1) The characteristics of CVD deaths were significant and of spatial clustering.

Consistent with the conclusions of previous international statistical studies, the CVD deaths were characterized as follows: male deaths were more than females; the mortality is higher in those of higher age; most of them are unemployed; and cardiocerebral infarction was the main cause of death. In addition, the deaths were mainly distributed in the old urban area, and there was a significant spatial clustering phenomenon [34]. Previous studies have shown that alcohol abuse [35] and smoking [36] can affect the cardiovascular system, and males are more likely to do alcohol abuse and smoking than females, causing the difference in mortality. The risk of cardiac arteriosclerosis will increase with age; that is, the mortality is higher in those of higher age [37]. Working can reduce the risk of CVD [8], and the risk of CVD is lower in residents engaged in manual work, but higher in those unemployed. Low temperature and sudden drop in air temperature will
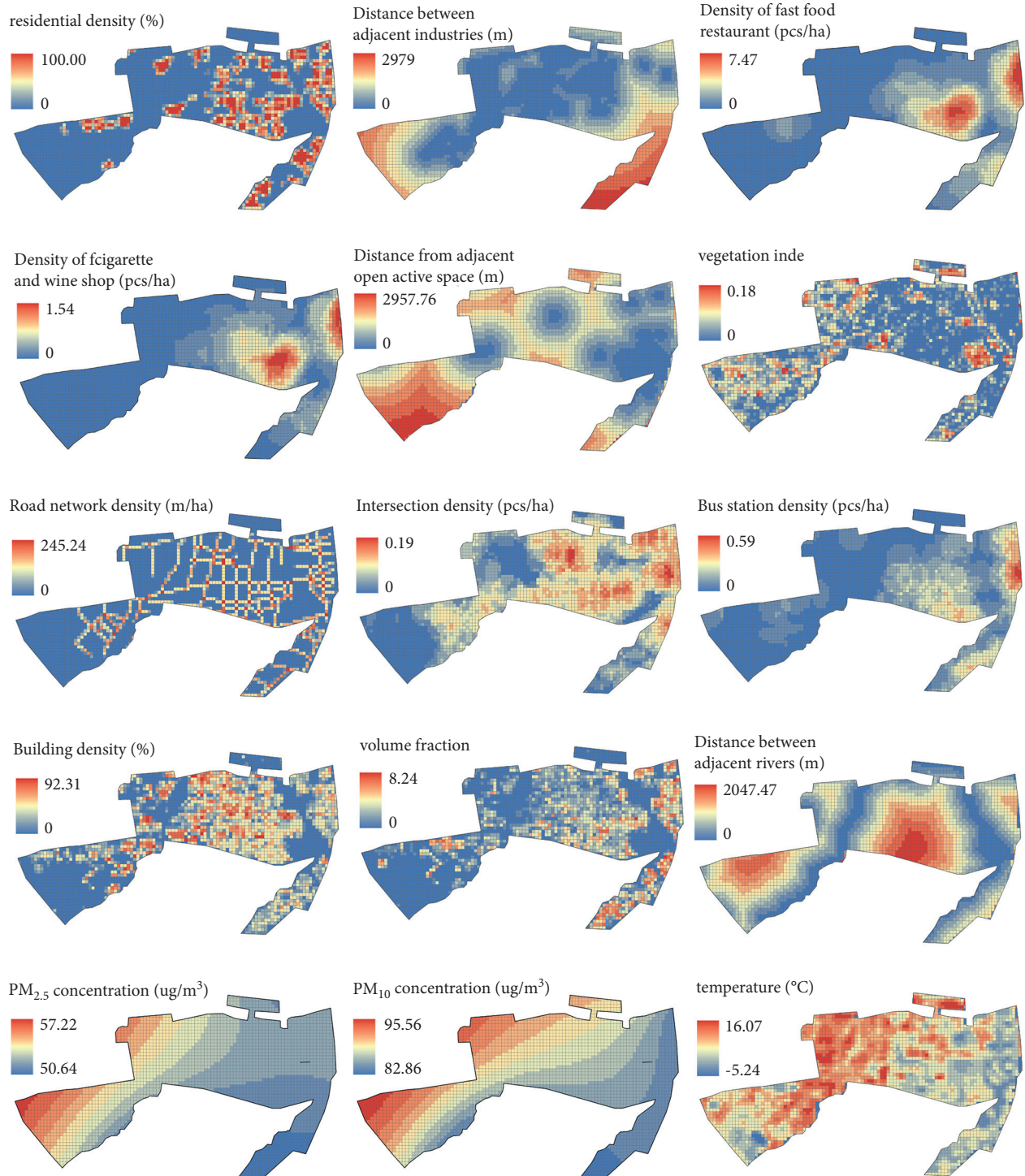
FIGURE 6: Spatial distribution of environmental influencing factors.

cause vasoconstriction and elevated blood pressure of human body, increasing the risk of sudden myocardial infarction [23, 24]; therefore, autumn and winter are the seasons of high incidence of CVD, and the mortality is higher in the urban areas with lower urban temperature [38].

(2) The higher density of restaurants and cigarette and wine shops will increase the CVD mortality.

It has been further confirmed in this study that the density of restaurants and smoking restaurants around the residential area showed a significant positive correlation with the CVD mortality. The daily activity distance that can meet the basic material and life needs of residents by 10-min walking is 500 m [39]. Handy et al. [40] found that the frequency of walking to businesses was positively

correlated with the number of businesses within the service radius; therefore, people in the areas with highly dense restaurants and cigarette and wine shop will have higher probability of dining out and buying cigarettes and wines. In the study area, located in northern Anhui, the restaurant diet contains high fat and salt, and the dense dining facilities will bring more cooking fumes, sewage, and other environmental pollution, in which a lot of fat and cholesterol oxide [17] will lead to hypertension, hyperlipidemia, and other metabolic disorder CVD, and the increased probability of buying alcohol and tobacco will raise the risk of alcoholism, thereby causing sudden death [17].

(3) The influence of environmental factors related to particulate matter concentration on the CVD mortality.

Consistent with the conclusion of previous studies abroad, long-term exposure to high levels of pollutants will affect the human circulatory system [41], thereby inducing CVD mortality. According to the data of disease burden from Barcelona, Spain, it was estimated that 849 cases of CVD could be prevented each year if the exposure level of air pollution could be reduced to the WHO recommended level [42]. There is a higher risk of death for residents in areas with high concentrations of particulate matter. The finding is further confirmed in this study: the increased traffic volume [43], dense bus stops [44], and dense residential and spatial forms [45, 46], low green space coverage [47], and longer distance from rivers [47] can increase the CVD mortality, of which the possible reason may be associated with the increased concentration of local particulate matter and other noise pollution and psychological harm.

(4) Discussion on the correlation between industrial and open activity space and CVD mortality.

Different from general studies, it is found that there is no significant relationship between CVD mortality and industry in this study, for which reason may lie in that northeasterly wind is the most prevalent in the study area throughout the year, and there are few residential areas located downwind of industry, only accounting for less than 10% of the study units; therefore, the influence of industry on the study unit is not significant. Meanwhile, this study also reveals that CVD mortality is lower in the areas farther from open space area, and the possible reason is that the study area is an old urban area that lacks green space, there are only 3 parks with the area of more than 1 ha within or 500 m from the study area (aggregated distribution), and there is relatively little activity space, which also is the common fault of an old urban area. The elderly are at high risk of dying from CVD, and they will deliberately choose the area around the open activity space when choosing housing, which will have a certain impact on the conclusion.

5. Conclusions and Prospects

To explore the distribution characteristics of cardiovascular disease (CVD) in urban residents and the environmental influencing factors, this study analyzed the basic characteristics and spatial distribution characteristics of CVD death population, explored the correlation between multiple environmental factors and CVD mortality, and revealed the factors affecting the risk of death and its mechanism, with the main conclusions as follows:

(1) The correlation study conducted in this study is of good scientific nature and research significance, and the conclusions are consistent with international studies. Therefore, the study methods can provide some research ideas for the construction of healthy cities in the Huaihe River Basin or other regions.

(2) With obvious characteristics, CVD deaths are mainly distributed in the vicinity of Unit YH09 in Zhanggongshan Park and the north of Unit YH08 near the Huaihe River. Therefore, attention should be paid to the key areas with the maximum number of CVD deaths and the population at high risk, and medical input should be increased to reduce the death risk.

(3) Because long-term exposure to high pollutants in the environment will cause CVD death, the following measures can be taken to reduce the CVD mortality in the construction of healthy cities in the future: increase traffic arterial roads around residential areas to reduce the traffic volume and vehicle exhaust emissions in residential areas; increase open space to reduce living and spatial density; plant more plants and building more rivers to reduce the local air pollution; and make dispersed layout of commercial facilities, such as restaurants and cigarette and wine shops, around residential areas to effectively reduce the risk of CVD death. Meanwhile, attention should also be paid to the influence of the areas with a higher density of the elderly on the analysis results.

Appropriate urban planning and policies can improve the active health interventions in cities and reduce the CVD mortality. To a certain extent, it can also provide more research perspectives for the construction of healthy cities in the future, and promote the development of healthy cities. In addition, this study is based on panel data, and little consideration is given to the changes of built environment and CVD mortality caused by time changes, so time series related research will be carried out in the future. At the same time, different lifestyles will affect CVD mortality, so further research on the correlation between residents' lifestyles and CVD mortality will be conducted in the future.

Data Availability

The datasets used during this study are available from the corresponding author on reasonable request.

Disclosure

Conflicts of Interest

The authors declared that they have no conflicts of interest in this work.

Acknowledgments

References

[1] H. Beltránsánchez, S. Soneji, and E. M. Crimmins, "Past, present, and future of healthy life expectancy," *cold spring harbor perspectives in medicine*, vol. 5, no. 11, 2015.

[2] H. H. A. Kyu, D. Abate, K. H. Abay et al., "Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017," *The Lancet*, vol. 392, pp. 1859–1922, Article ID 10159, 2018.

[3] S. S. A. S. Jiaying Zhao, "Timing of urbanisation and cardiovascular risks in Thailand: evidence from 51 936 members of the Thai cohort study, 2005–2009," *Journal of Epidemiology*, vol. 24, no. 6, pp. 484–493, 2014.

[4] S. Rani and W. Anthony, "Current levels of salt knowledge: a review of the literature," *Nutrients*, vol. 6, no. 12, pp. 5534–5559, 2014.

[5] Y. Xu and F. Wang, "Built environment and obesity by urbanicity in the US," *Health & Place*, vol. 34, pp. 19–29, 2015.

[6] D. W. Barnett, A. Barnett, A. Nathan, J. V. Cauwenberg, A. Barnett, and au fnm, "Built environmental correlates of older adults' total physical activity and walking: a systematic review and meta-analysis," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, 2017.

[7] M. Stevenson, J. Thompson, S. á. De et al., "Land use, transport, and population health: estimating the health benefits of compact cities," *Lancet*, vol. 388, pp. 2925–2935, 2016.

[8] R. Ewing, G. Meakins, S. Hamidi, and A. C. Nelson, "Relationship between urban sprawl and physical activity, obesity, and morbidity – update and refinement," *Health & Place*, vol. 26, pp. 118–126, 2014.

[9] F. Ballester, J. Tenias, and M. Perez-Hoyos, "Air pollution and emergency hospital admissions for cardiovascular diseases in Valencia, Spain," *Journal of Epidemiology & Community Health*, vol. 55, pp. 57–65, 2001.

[10] Y. Guo, J. P. Teixeira, and N. Ryti, "Ambient particulate air pollution and daily mortality in 652 cities," *New England Journal of Medicine*, vol. 381, no. 8, pp. 705–715, 2019.

[11] M. E. Héroux, H. R. Anderson, R. Atkinson, B. Brunekreef, and H. Walton, "Quantifying the health impacts of ambient air pollutants: recommendations of a WHO/Europe project," *International Journal of Public Health*, vol. 60, no. 5, 2015.

[12] M. Oh, K. Shin, K. Kim, and J. Shin, "Influence of noise exposure on cardiocerebrovascular disease in Korea," *The Science of the Total Environment*, vol. 651, pp. 1867–1876, 2018.

[13] R. D. Brook, S. Rajagopalan, C. A. Pope, J. R. Brook, and J. D. Kaufman, "Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association," *Circulation*, vol. 121, no. 21, pp. 2331–2378, 2010.

[14] C. Hong-han, C. Hong-wei, H. F. Jiang-tao, L. Zhao-li, S. Bing, and S. Jing, "Health-based risk assessment of contaminated sites:Principles and methods," *Earth Science Frontiers*, vol. 13, no. 001, pp. 216–223, 2006.

[15] W. Lan, J. Xiji, S. Wenyao, Z. Xiaojing, and T. Jian, "Impact of urban built environment on respiratory health and its planning strategy: a case study of a district in shanghai," *China City Planning Review*, vol. 42, no. 06, pp. 15–22, 2018.

[16] S. S. Mehraj, G. A. Bhat, H. M. Balkhi, and T. Gul, "Health Risks for Population Living in the Neighborhood of a Cement Factory," *Academic Journals*, vol. 7, no. 12, pp. 1044–1052, 2013.

[17] T. D. K. Y. Fei-hong, "Analysis on oil fume particles in catering industry cooking emission," *Environmental science*, vol. 33, no. 006, pp. 1958–1963, 2012.

[18] P. Malambo, A. P. Kengne, A. De Villiers, E. V. Lambert, and T. Puoane, "Built environment, selected risk factors and major cardiovascular disease outcomes: a systematic review," *PLoS One*, vol. 11, no. 11, 2017.

[19] W. Babisch, "Road traffic noise and cardiovascular risk," *Noise and Health*, vol. 10, no. 38, pp. 27–33, 2008.

[20] W. D. Douglas, "Epidemiologic evidence of cardiovascular effects of particulate air pollution," *Environmental Health Perspectives*, vol. 109, pp. 483–486, 2001.

[21] M. Furuuchi, T. Murase, M. Yamashita et al., "Temperature distribution and air pollution in phnom penh, Cambodia-influence of land use and the mekong and tonle sap rivers," *Aerosol and Air Quality Research*, vol. 6, no. 2, pp. 134–149, 2006.

[22] S. Valdés, C. M. Araque, F. G. Torres et al., "Ambient temperature and prevalence of obesity in the Spanish population: the D i@bet.es study," *Obesity*, vol. 22, no. 11, 2014.

[23] K. Bhaskaran, S. Hajat, A. Haines, E. Herrett, P. Wilkinson, and L. Smeeth, "Short term effects of temperature on risk of myocardial infarction in England and Wales: daily time series regression analysis of the Myocardial Ischaemia National Audit Project (MINAP) registry," *BMJ British Medical Journal*, vol. 341, no. 7768, 2010.

[24] J. A. Patz, D. Campbell-Lendrum, T. Holloway, and J. A. Foley, "Impact of regional climate change on human health. Nature 438: 310-317 (17 november)," *Nature*, vol. 438, no. 7066, pp. 310–317, 2005.

[25] Y. Xu and L. Wang, "GIS-based analysis of obesity and the built environment in the US," *Cartography and Geographic Information Science*, vol. 42, no. 1, pp. 9–21, 2015.

[26] H. M. GriffithsC, G. M. A. Jordan, and H. Saunders, "Neighbourhood typologies and associations with body mass index and obesity: a cross-sectional study," *Preventive Medicine: An International Journal Devoted to Practice and Theory*, vol. 111, pp. 351–357, 2018.

[27] P. Zhang, J. Lee, G. Kang et al., "Disparity of nitrate and nitrite in vivo in cancer villages as compared to other areas in Huai River Basin, China," *The Science of the Total Environment*, vol. 612, pp. 966–974, 2018.

[28] Y. Juna, "Research on the spatial pattern of population destribution in the Huaihe River basin," *Journal of Xinyang Normal University*, vol. 30, no. 1, pp. 525–528, 2014.

[29] T. W. Hwang and F. B. Wang, "Dynamics OF a dengue fever transmission model with crowding effect IN human

population and spatial variation," *Discrete and Continuous Dynamical Systems - Series B*, vol. 1, no. 1, pp. 147–161, 2013.

[30] C. Varga, D. L. Pearl, S. A. Mcewen, J. M. Sargeant, F. Pollari, and M. T. Guerin, "Area-level global and local clustering of human Salmonella Enteritidis infection rates in the city of Toronto, Canada, 2007–2009," *BMC Infectious Diseases*, vol. 15, no. 1, p. 359, 2015.

[31] G. A. Colditz, W. C. Willett, M. J. Stampfer et al., "The influence of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire," *International Journal of Epidemiology*, p. 392, 1987.

[32] W. Xiaoyanl and L. Meizhou, "The relationship of rank correlation coefficient and spearman rank correlation coefficient," *Journal of Guangdong Industry Technical Colleg*, vol. 5, no. 004, pp. 26-27, 2006.

[33] D. Khosrow, "Density estimation for statistics and data analysis," *Technometrics*, vol. 29, no. 4, p. 495, 2012.

[34] C. T. Y. Y. Liu Jie, "Exploratory spatial data analysis of patients with acute myocardial infarction from 2007 to 2014 year in Yangpu District, Shanghai," *Chin J Evid Based Cardiovasc Med*, vol. 8, no. 4, pp. 464–467, 2016.

[35] X. H. Ning, W. A. C. Cuan, and G. Hong, "Further research into relationship between alcohol consumption and plasma homocysteine," *Adv Cardiovasc Dis*, vol. 31, no. 03, pp. 445–448, 2010.

[36] S. Qian, Z. Nanbo, Y. Canqing, G. Yu, B. Zheng, and L. J. Chen Zhengming, "Sex-specific associations between tobacco smoking and risk of cardiovascular diseases in Chinese adult," *Chinese Journal of Epidemiology*, vol. 39, no. 001, pp. 8–15, 2018.

[37] L. Liu, J. G. Wang, L. Gong, G. Liu, and J. A. Staessen, "Comparison of active treatment and placebo in older Chinese patients with isolated systolic hypertension: Systolic Hypertension in China (Syst-China) Collaborative Group," *Journal of Hypertension*, vol. 16, no. 12, p. 1823, 1998.

[38] L. Q. Z. Z. Liu Xue-en, "Influence of temperature and cold air on mortality of cardio -and cerebral vascular diseases," *CHINA PUBLIC HE ALTH*, vol. 18, no. 8, 2002.

[39] L. Meng, "The planning strategies of a 15- minute community life circle based on behaviors of residents," *Urban Planning Forum01)*, vol. 1, pp. 111–118, 2017.

[40] S. Handy, X. Cao, and P. L. Mokhtarian, "Self-selection in the relationship between the built environment and walking: empirical evidence from northern California," *Journal of the American Planning Association*, vol. 72, no. 1, pp. 55–74, 2006.

[41] D. F. M. M. Chuan, "Association between ambient PM 10 / PM 2.5 levels and population mortality of circulatory diseases: a case-crossover study in Beijing," *Journal of Peking University*, vol. 45, no. 03, pp. 398–404, 2013.

[42] N. Mueller, D. Rojas-Rueda, X. Basagaa et al., "Health impacts related to urban and transport planning: a burden of disease assessment," *Environment International*, vol. 107, pp. 243–257, 2017.

[43] R. Sinharay, J. Gong, B. Barratt et al., "Respiratory and cardiovascular responses to walking down a traffic-polluted road compared with walking in a traffic-free area in participants aged 60 years and older with chronic lung or heart disease and age-matched healthy controls: a randomised, crossov," *Lancet (London, England)*, vol. 391, pp. 339–349, Article ID 10118, 2020.

[44] L. Zhao, X. Wang, Q. He et al., "Exposure to hazardous volatile organic compounds, PM10 and CO while walking along streets in urban Guangzhou, China," *Atmospheric Environment*, vol. 38, no. 36, pp. 6177–6184, 2004.

[45] L. Schweitzer and J. Zhou, "Neighborhood air quality, respiratory health, and vulnerable populations in compact and sprawled regions," *Journal of the American Planning Association*, vol. 76, no. 3, pp. 363–371, 2010.

[46] B. Tong, H. Lijian, H. Liang et al., "Impact of urban air pollution on surrounding areas in Beijing , Tianjin , and Hebei Provinces," *Acta Ecologica Sinica*, vol. 038, no. 012, pp. 4268–4275, 2018.

[47] D. Wang, K. L. Lau, R. Yu, S. Wong, T. Kwok, and J. Woo, "Neighbouring green space and mortality in community-dwelling elderly Hong Kong Chinese: a cohort study," *BMJ Open*, vol. 7, no. 7, 2017.

Research Article

Meta-Analysis of the Clinical Effect of MIS-TLF Surgery in the Treatment of Minimally Invasive Surgery of the Orthopaedic Spine

Wanliang Yang,[1] Xin Pan [ID],[2] and Xun Xiao[1]

[1]*College of Clinical Medicine, Shandong University, Jinan 250000, Shandong, China*
[2]*Qilu Medical College of Shandong University, Jinan 250000, Shandong, China*

Correspondence should be addressed to Xin Pan; 202036073@mail.sdu.edu.cn

Minimally invasive surgery (MIS) has already had a significant impact on surgical treatment (spine). Because they are less invasive, minimally invasive treatments are often preferred over open spine surgery. MIS and open spine surgery in terms of posterior lumbar fusion (PLF), lumbar disc herniation (LDH), and cervical disc herniation (CDH) were all observational studies based on randomized controlled trials. Seventeen RCTs and six observational studies were conducted. Chemotherapy had no effect on the long-term alleviation of the neck or arm pain in patients with CDH. In LDH, MIS was superior in terms of pain relief, rehospitalization rates, and improvement in quality of life. At the expense of increased perioperative endoscopic, readmission, and revision rates, MIS achieved a significant reduction in 2-year expenditures, fewer medical problems, and improved Oswestry score ratings. There is no evidence to support the use of MIS over open surgery for lumbar or cervical process disc herniation. In comparison, MIS-TLIF has several advantages, in addition to lower revision/readmission rates. However, MIS significantly increases the surgeon's radiation exposure, regardless of the patient's sign. However, the effect on patients is unknown. These findings could help patients make better decisions when comparing open spine surgery to minimally invasive spine surgery, especially given how much advertising is out there for MIS.

1. Introduction

Historically, "open surgery" has been the method of choice for spine surgery. A lengthy incision must be made in the area to be operated on for the surgeon to be able to examine and analyze the anatomical structures. More back and neck disorders may now be treated with minimally invasive surgery thanks to recent technical advances. Because MISS does not require a long incision, it reduces the risk of substantial damage to the muscles around the spine. Pain and recovery time are reduced. The indications for minimally invasive spine surgery are similar to those for open surgery. Spine surgery is usually shown only after nonsurgical treatments like drugs and physical therapy have failed to relieve severe symptoms. Also, surgery is only shown if your doctor has discovered the specific cause of your pain, such as a herniated disc or spinal stenosis. There are a variety

of minimally invasive techniques available. Every one of them has one thing in common: reduced muscle injury and smaller incision size. Using minimally invasive procedures to undertake common operations such as lumbar decompression and spinal fusion, which are both effective, decompression lowers pressure on the spinal nerves by removing bone fragments or a herniated disc. This procedure is used to correct problems with the spine's small bones (vertebrae). The goal is to heal the painful vertebrae and then fuse them together to form a single strong bone structure. The goal of this article is to explore spinal decompression and fusion techniques that are minimally invasive. MISS (minimally invasive spine surgery) is another term for this procedure. Using specialized tools, surgeons can reach the spine through small incisions during these operations. During open surgery, the doctor makes a six-inch incision or less and then moves the muscles to the side to see the spine.

The surgeon can now reach the spine and remove sick or injured bone or intervertebral discs using the side muscles. Additionally, the surgeon can see right away where cages, bone grafts, and screws are needed to supply support and ease healing of the spinal bones throughout the procedure. Open surgery may cause a significant level of criticality because of the possibility of the muscle being harmed during the tugging or "retraction" procedure. While muscle retraction is designed to aid the surgeon in finding the location of the problem, it often affects anatomy more than the surgeon needs. Therefore, muscle injury is more likely to occur because of this, and patients may feel postoperative pain that is distinct from the back discomfort they were experiencing before the surgery. It is possible that this will result in an extended recovery time. Added complications may include higher blood loss and infection risk because of the larger incision and harm to soft tissues. This technique was created to treat spinal illnesses while minimizing injury to muscles and other natural spine systems. Additionally, it allows the surgeon to see only the part of the spine that is affected by the problem (the problem area). In addition, MIS offers the advantage of requiring fewer incisions, resulting in less bleeding and requiring fewer hospital stays. The most common technique is the tubular retractor. A tiny incision is made and a tubular retractor is inserted into the spinal column through the skin and soft tissues. This builds a tunnel to the affected area of the spine. The tubular retractor keeps the muscles relaxed throughout the process. The surgeon gets access to the spine using small devices that are integrated into the tubular retractor's central core. To install fusion devices such as screws or rods, the retractor must first be removed from the area of the injured bone or disc. For some surgeries, many retractors or incisions may be needed.

The surgeon uses fluoroscopy to guide him through the process of creating the incision and inserting the retractor. During the surgery, a screen shows the surgeon real-time X-ray pictures of the patient's spine. During surgery, the surgeon will typically use a microscope to examine the spine's key components, such as the discs and joints. After surgery, the tubular retractor is withdrawn and the muscles are allowed to heal. The likelihood of muscular injury is more common with open surgery; in the spine surgery sector, new minimally invasive procedures are constantly being developed. An endoscope, for example, is currently being used by some spine surgeons to gain access to a problem site in the spine. Anesthesia for MISS is most administered as either general anesthesia (during which you stay unconscious) or regional anesthesia (during which you are awake during the procedure).

1.1. Overview of the Manuscript. In this manuscript, the meta-analysis of the clinical effect of MIS-TLIF surgery in the treatment of minimally invasive surgery of the orthopaedic spine and the common comparison between the MIS and the TFIL are discussed in the following section, which is followed by the literature survey. Participation in the research is subject to certain requirements. Extraction of information, the quality of the research, and analytical statistics are used to analyze data. Evidence levels are used to categorize the effects of research investigations, and patients and the process of choosing the method are discussed as follows.

2. Minimally Invasive Spine Surgery Is a Common Surgery

The compression of a nerve by a herniated disc in the lower back can cause severe leg discomfort, numbness, and even paralysis in certain cases. To ease these symptoms, the disc is surgically removed. A diskectomy is the medical term for this procedure.

To remove the herniated disc, the patient is turned face down and a tiny incision is made. During the surgery, a small part of the lamina bone is removed. This gives a better view and an idea for the surgeon to visualize the spinal nerve and the disc. A surgeon (doctor) delicately retracts the nerve and removes only the affected disc, leaving the rest of the nerve intact. It is also possible to employ this minimally invasive method to repair herniated cervical discs in the neck. The technique is known as MIS posterior cervical foraminotomy/discectomy.

Open lumbar fusions may be done via the back, belly button, or side. Lumbar fusions may also be performed minimally invasively. The surgeon can approach the spine from the side, reducing the amount of spinal nerve displaced during the surgery enabled by the transforaminal lumbar interbody fusion (TLIF). After the patient is placed face down during a MIS-TLIF, the surgeon installs one retractor on each side of the spine, completing the procedure. This avoids rupturing the midline ligaments and bone. The surgeon removes the lamina and disc, then places the bone transplant into the disc region, fastening it with screws or rods as needed. To boost the likelihood of healing, the surgeon may choose to use a bone transplant in addition to the patient's own bone on occasion. Aside from being performed minimally invasively from the side, spinal fusions are also performed regularly. The side approach is enabled by the Direct Lateral Interbody Fusion (DLIF) and Extreme Lateral Interbody Fusion (XLIF). They are superior to conventional spinal fusion procedures in that they do not injure the back muscles or strain or pull on the spinal canal nerves.

An oblique lateral interbody fusion is a more contemporary variation of this surgery that is performed on the left side of the body (OLIF). OLIF, like XLIF and DLIF, needs the use of a side incision. Instead, OLIF enters an oblique position, avoiding the psoas muscle (the muscle on the side of the spine). All three of these lateral strategies produce outcomes that are comparable. MISS, like any other operation, entails some inherent hazards that must be understood and accepted. Some studies show that MISS has similar issues to those associated with open spinal fusion surgery, but some studies show that MISS has a lower infection rate than open spinal fusion surgery. Before surgery, your doctor will go over each risk with you in detail, and they will take specific precautions to help avoid any negative outcomes.

3. The Following Issues May Arise Because of MISS

Antibiotics are given to the patient before, during, and often after the surgery to lower the risk of developing an infection. It is normal to experience some bleeding after surgery, but it is usually not significant. There is a pain in the area where the graft was placed. Occasionally, a small number of people have persistent discomfort at the location of their bone transplant. Some symptoms can recur on a regular basis. Some people may revert to their first symptoms. Pseudarthrosis, insufficient bone growth, causes incomplete spinal fusion healing. If this happens, further surgery may be necessary to produce a solid bone union. Pseudarthrosis is more frequent among smokers, who are more susceptible. Nerve injury is a big issue. These operations may injure nerves or blood vessels. These kinds of ramifications are quite uncommon. Blood clots are formed. It is also possible that blood clots will form in the legs because of surgery, which is unusual. If they shatter and travel to the lungs, they are a major hazard to the patient's health. Minimally invasive procedures can result in shorter hospital stays since they are less intrusive. MISS patients typically return home on the same day or within one to two days of being admitted to the hospital. The exact length of their hospital stay varies depending on their condition and therapy. Most patients who undergo conventional surgery are admitted to the hospital for three to five days after the procedure.

Because minimally invasive therapies do not damage muscles or soft tissues, postoperative discomfort is believed to be less severe than that experienced following traditional open procedures in most cases. While you can expect some discomfort, current advances in pain management help your doctor manage and treat your pain. Physical therapy may be recommended by your doctor to aid you in regaining strength and speeding up your rehabilitation. Depending on the procedure and your overall health, this will vary. Specific workouts might aid you in regaining the strength you need to return to work and your normal activities after an injury. It might take months for the bone to harden after a fusion. On the other hand, your degree of comfort will often improve far more quickly. For the fused spine to heal properly, it is vital that it keeps proper alignment throughout the healing process. On the job, you will learn to move appropriately, reposition, sit, stand, and walk. The amount of time it will take you to return to your typical activities after a MISS varies based on your unique operation and condition. In the days after surgery, your doctor will check on you to ensure you are recovering well.

Medical and surgical therapies are evaluated objectively through research. The quality of evidence decides how much research impacts management. Class I data from prospective randomized controlled trials (RCTs) are the most persuasive in showing a treatment's effectiveness. In this research, the following conditions are analyzed, such as [1]. Five studies comparing MIS with open surgery for CDH were chosen as Class I trials. Four RCTs and systematic analysis were included. The RCTs included 200+ MIS patients and 150+ open surgery patients, following them for more than 100 weeks (about 2 years) [2]. Only one of the four RCTs (8 percent) used a discectomy without fusion on 19 of 200+ open surgery participants. The remaining 150+ patients had anterior cervical discectomy. Only 37 out of 200+ MIS patients (15%) had fusion, while 100 out of 150+ had posterior foraminotomy. A posterior cervical discectomy was performed on 60 individuals, and an amniotomy or discectomy was performed on 18 patients (specific procedure per patient not reported). In RCTs, only 95 of 200+ MIS patients (42%) employed an anterior approach, compared to 190 of 200+ conventional open patients (88%) [3, 4].

This review shows that MIS did not enhance neck pain and arm discomfort, compared to TLIF for CDH. While MIS helped relieve temporary neck discomfort, when lumbar patients were included in the pooled estimate analysis, this improvement was not statistically significant. Minimally invasive vs. open surgery Twelve trials comparing MIS with treatment of LDH produced Class I evidence. It includes ten RCTs and two systematic reviews. In the 10 RCTs, 489 individuals receiving MIS and 500+ patients with traditional open surgery were recruited in the 10 RCTs. Both groups did not obtain fusion. Out of 500+ open patients, 500+ got discectomy, and 19 had percutaneous nucleosome, with the follow-up durations of 50–100 weeks (about 2 years) [4–6].

On average, MIS is inferior to open LDH surgery for leg discomfort, less back pain, lifestyle and quality of life, and rehospitalization. Conversely, MIS was linked to a decreased risk of infection and a reduced hospitalized duration. Postoperative Oswestry Disability Index (ODI) scores were not altered. The authors suggest that any alleged decreased infection incidence is linked with MIS. When compared to open surgery, MIS exposed the surgeon's thyroid and eye more than tenfold, his chest 14-fold, and his hand 20-twofold. Surgery with a minimum of invasiveness against surgery with a maximum of invasiveness: disc herniation is a condition in which the disc protrudes from its normal location (cervical or lumbar) [5, 6]. Data from 14 RCTs on CDH and LDH were combined to form a single comprehensive assessment. When compared to traditional open surgery, the study discovered that minimally invasive surgery (MIS) decreased infection rates while increasing nerve root damage, durotomy, and reoperation rates. However, none of these changes were statistically significant. At the rear of the neck, the lumbar spine is fused [7].

4. Literature Survey

Lumbar fusion surgery treats spinal instability, stenosis, spondylolisthesis, and degenerative disc disease symptoms, normalizing motion and stability while keeping load-bearing ability and aligning the system. Lumbar fusion operations

have grown in popularity in the US during the last several decades [8].

Transitioning from traditional open surgery to less invasive approaches is the method of choice for many surgical procedures. According to postoperative histology and imaging examinations, standard open procedures leave traces of scar during tissue development, muscle retraction, and stripping, all of which have a negative impact on outcomes and increase the likelihood of reoperation. To help maintain paraspinal muscular anatomy and bone architecture, minimally invasive treatments employ a muscle-dilating approach. These treatments have been proven to reduce the incidence of iatrogenic soft-tissue injuries [9] [6–8]. Reduced postoperative discomfort, decreased intraoperative blood loss, a shorter postoperative staying in hospital, fast recovery to normal activities, and a lower reoperation rate are just a few of the reasons that minimally invasive spine (MIS) therapies are becoming increasingly popular. Due to the lack of long-term data on patients receiving minimally invasive spinal fusion for severe back pain, the use of these procedures instead of standard open fusion methods continues to be controversial in the medical community. We want to add to the data in this discussion by supplying long-term, prospectively recorded results from one of the biggest presently known series of MIS-TLIF by following patients up to 28 months (about 2 and a half years) [10]. According to our findings, MIS-TLIF beat open TLIF in all domains, except for the fusion rate. As reported in the meta-analysis, the MIS-TLIF procedure had an identical fusion rate as the open TLIF procedure but with a shorter hospital stay, faster ambulation, and less blood loss [11]. Staying in the hospital for less time also reduces medical expenditures because MIS-TLIF enhances ambulation speed [12]. The MIS group's postoperative VAS and ODI scores were lower than the open groups. Although the preoperative clinical and functional baseline measures were equal across the groups, the open group's postoperative ratings were lower than those of the MIS groups. So, MIS-TLIF improved outcomes while reducing trauma [13].

In recent years, TLIF has gained widespread acceptance and recognition as a surgical method that significantly reduces the likelihood of relative nervous system disorders, despite the advancement of fusion technology. However, the open TLIF technique requires paraspinal muscle splitting to be successful. It does not hurt, but it ruptures a large piece of the posterior compartment, causing muscle stiffness and low back discomfort. As a result of its advantages over standard open surgery, such as less harm to spinal soft tissues and paravertebral muscles, it has gained popularity in recent years. The advantages of MIS-TLIF can be linked to the fact that less paravertebral muscle dissection and retraction occur throughout the procedure [14].

According to the study, it is found that the MIS-TLIF and open TLIF have no difference in the surgery time, complication rates, or reoperation rates. Despite this, the compilation and reoperation rates for the MIS are better than the open TLIF process. One of the most persuasive arguments is that the learning curve for minimally invasive surgical abilities is steep, requiring added years of experience

to perfect the necessary skills [15]. The advancement of surgical devices and technology may pave the way for realizing the other critical goals. Our meta-analysis has several flaws that need to be addressed. Because we only looked at prospective and retrospective research, issues including data bias, nonblinding or improper, inadequate baseline comparisons, and data collection must be addressed. For the second time, the current study relied on a small number of outcome indicators; clinical outcomes should be assessed by a variety of alternative aims or subjective characteristics. In addition, there were inconsistencies in the definition and assessment of fusion during the study. A further point to mention is that many of the investigations were completed in a short or medium amount of time. More long-term follow-up studies must be reviewed to analyze the evaluation process for the effectiveness of these measures. Finally, an inherent bias was added to the data by combining them all together [16]. Data collection and inappropriate or nonblinding were challenges that arose due to solely looking at prospective and retrospective studies. The use of MIS-TLIF is also not related to an increase in complications or reoperations. Comparing open surgery and MIS, there are fewer lesions, better outcomes, and the same fusion rate. Surprisingly, MIS-TLIF has been linked to fewer complications and reoperations. This may be due to increased acceptability of MIS-TLIF, familiarity with the surgery, ability in MIS abilities, and development of proper surgical instructions and instruments. Added high-quality research is needed to confirm and compare these findings [15, 16]. We searched the CNKI, Wanfang, and VIP databases for relevant information using the keywords "MIS" and "TLIF". Many researchers recommend restricting the search to English-language articles. The retrieval period ran from the database's creation until January of the current year. It was also searched for articles related to the original research and review papers in the reference lists of those articles. Two reviewers independently reviewed all publications based on their titles, abstracts, or full texts. The two reviewers assigned to them knew the writers and their journals [17]. On a particular research project, diverse viewpoints were heatedly debated until an agreement was set up.

4.1. Participation in the Research Is Subject to Certain Requirements.
The criteria for consideration were as follows: (3) at least one clinic result or perioperative data were presented in the article; (4) the patients had degenerative disc disorders (disc herniation, canal stenosis, or spondylolisthesis); this study excludes the patient who underwent open TLIF or MIS-TLIF for other illnesses. These patients were likewise eliminated from the study [18, 19].

4.2. Extraction of Information.
From the research that was included, we were able to extract information for the categories listed as follows: (1) the names of the authors and the year of publication; (2) the study strategy; (3) the type of evidence; (4) the total number of patients recruited, as well as the number of patients in each group (MIS-TLIF versus open TLIF); (5) the average amount of time spent following

up; (6) the average follow-up time and rate (in percentage points); (7) the patients' average age; (8) the proportion of patients who are male and female; (9) predictive diseases (like degenerative disc disease, spondylolisthesis, and others are classified into three categories); (10) clinical outcomes; (11) perioperative data points; (12) inclusion/exclusion criteria; (13) information about the postoperative period; (14) the total number of lumbar segments treated; (15) the definition of fusion and the assessment of the measure; (16) the types and quantity of grafts used; (17) the number and type of cages used; (18) the screw fixation technique and the number of issues met; (19) the number of complications met. We had not predicted any problems ahead of time. Moreover, the total number of challenges was tallied up for analysis [19, 20].

4.3. The Quality of the Research. Levels of Evidence (2009) Class I claims were supported by high-quality randomized controlled studies (RCTs). Class II evidence includes RCTs of moderate to inferior quality and correct cohort studies. Class III evidence includes cohort and case-control studies of moderate or inferior quality. The Case Series Study was in Class IV Evidence. Two reviewers separately evaluated the publications, and disagreements were discussed until a consensus was reached [21, 22].

4.4. Analytical Statistics Are a Type of Statistics That Is Used to Analyze Data. WMDs were used to measure the ODI, average blood loss, hospitalized duration, time taken for the surgery, and VAS. The relative risk (RR) measure and 95% confidence interval were used for dichotomous data. Each study has its own unique set of criteria for evaluating the efficacy of the fusion procedure. Consequently, the total number of studies within each category was not equal. In this research, a random-effects model was preferred over a fixed-effects model because it supports a distribution that the fixed-effects model does not. All tests required a p-value of 0.05 or less. The publication bias was investigated using funnel plots. Asymmetry suggests a publication's prejudice, while symmetry implies no bias. We could conduct added database analysis using Review Manager [5].

4.5. Evidence Levels Are Used to Categorize the Effects of Research Investigations. Randomized double-blind placebo-controlled trials are appointed as Class 1 trials (randomized outcome measures) [23]. Experiments were conducted without a randomization procedure (Class II). Unlike Class I, this group is not random. Case-control, cohort, and interrupted time studies processed with controls are all examples of Class III Observational Studies. Observational studies without controls are Class IV (Class III, without controls) and Ability (Class V) (invited only can comment).

5. Patients and the Process of Choosing the Method

Between 2003 and 2010, 318 MITLIF surgeries, as shown in Table 1, were conducted on 304 consecutive patients, all of whom were treated using a paramedian muscle-sparing technique. During the operation, 120 men and 184 women were included in the research. The mean age (range, 19–93) and gender (male) at surgery were 62.4 and 19.4, respectively. The senior author performed all 318 MITLIF surgeries on all 318 patients from an outpatient neurosurgery spine clinic. Following a thorough clinical history, physical exam, and lumbar spine radiological study, the diagnosis was made. When spondylolisthesis and retrolisthesis were diagnosed, they were classified according to the Meyerding classification (I–IV). All the patients were examined outside of a hospital environment.

Complicating matters, the patient also has several medical conditions, which are listed in the table. Spondylolisthesis (66%) was the most common clinical condition, followed by central spinal stenosis (47%) and foraminal stenosis (34%). The most common clinical findings were persistent severe low back pain, neurogenic claudication, and radiculopathy-related symptoms [5]. Patients experienced symptoms for an average of 68.6 months (about 5 and a half years) before surgery. Previously, 70 individuals had lumbar surgical operations performed on them (23%). When clinical and imaging data agreed, as well as symptoms that had lasted for more than six months and had not responded to nonoperative treatment, patients were considered candidates. The table shows patients who had conservative nonoperative therapy. If a neurological impairment was deteriorating or if the patient presented with considerable, incapacitating pain and a good correlation between clinical and radiographic data, surgical therapy was started no sooner than 6 months. If nonoperative treatment fails, surgical intervention is recommended as the last resort for patients who were physically unsuited for surgery, patients who had a recent bleeding diathesis, patients who were infected, patients who had incongruent clinical and radiographic data, and patients who were infected [5, 22, 23].

In the field of outcome measurement, outcome measures are a type of outcome measurement. Low back pain, back-related functional disability, and physical and mental quality of life were all improved when the visual analogue scale (VAS), the ODI, and the Short-Form 30+ were used to assess patients' reported results (SF-36). During the recruitment process, patients were requested to complete these validated questionnaires, and they were also asked to complete them again at various stages throughout the postoperative period. The presurgery scores, all later follow-up, and change scores were all analyzed for significance. A secondary outcome measure was the rate of fusion, the rate of reoperation, the amount of intraoperative blood loss, and the length of time spent in the hospital following the procedure. To assess fusion, radiologists separately and blindly reviewed radiological images taken from the lumbar anterior-posterior and lateral flexion/extension views during the postoperative phase. We considered fusion effective when there was no considerable motion or angulation at the fused level, no implant latency, no hardware loosening or breakage, and bridging bone growth.

TABLE 1: Recommendation for patient population.

Recommendation	Population of the patient
As compared to open surgery conducted normally, MIS	
1. It has no effect on short-term functional outcomes.	
2. It does not have a negative impact on long-term performance.	
3. It does not alleviate acute arm discomfort.	Cervical disc
4. It does not supply relief from persistent arm discomfort.	herniation
5. It supplies immediate relief from severe neck aches.	
6. Does not supply relief from chronic neck pain.	
When compared to conventional open surgery, MIS	
1. It does not improve function in the short term.	
2. It does not have a negative impact on long-term performance.	
3. It does not supply enough relief for leg pain.	
4. Low back pain is not well relieved by this medication.	
5. It has a greater likelihood of requiring rehospitalization.	
6. The quality of life is negatively affected by the procedure.	
7. The chance of surgical site and infection issues is reduced.	Lumbar disc
8. It is possible that it is associated with a shorter period of hospital stay.	herniation
9. It does not show Oswestry Disability Index scores after at least six months after surgery interns of statistically significant difference.	
10. In addition, it exposes the surgeon to more than ten times the amount of radiation that would otherwise be supplied to the thyroid or the eyes.	
11. Exposure to more than 14 times the quantity of radiation supplied to the chest is imposed on the surgeon.	
12. Subjects the surgeon's hand to more than 22 times the amount of radiation received by the patient.	
According to the existing practice of open surgery, as compared to open surgery as it is now practiced,	
1. The incidence of nerve root damage is growing.	
2. The incidence of nerve root injuries is decreasing.	
3. A rise in the number of inadvertent durotomies that occur	Disc herniation
4. Increased reoperation rates are becoming more common as a trend.	
5. Infection-prevention trends in the United States	
When comparing MIS with TLIF and open TLIF	
1. leads to much less blood loss than the latter.	
2. It leads to a considerable reduction in back pain on the second postoperative day after surgery.	
3. It involves much greater amounts of intraoperative radiation time.	
4. It does not need a major increase in overall operating time.	
5. Despite considerable improvement in the ODI over a brief period, there is no statistically meaningful difference in terms of long-term clinical outcome.	
6. There is no statistically significant difference in the radiographic outcome.	Posterior lumbar fusion
7. It has led to reducing the number of hospitalizations of the patients	
8. It has reduced the amount of time necessary to a normal life routine.	
9. It has reduced indirect expenses.	
10. During a two-year period, he has cut social expenses.	
11. It has resulted in increased narcotic independence	
When compared to open TLIF/PLIF, MIS-TLIF/PLIF has the following benefits over the former:	
1. A faster rate of revision.	
2. A higher risk of hospital readmission.	Posterior lumbar fusion
3. There has been no change in the incidence of surgical complications since the study began.	
4. A reduction in the number of medical problems	

The most persuasive evidence supporting MIS compared to open spine surgery in the patient group suffering from CDH, LDH, and PLF is reviewed.

6. Data Analysis Using Statistical Methods: Analyzing the Information

To compare the ODI, SF-36, and VAS values for each time, Student's t-test was used. Consideration is decided based on the p-values, i.e., p-values <0.05 are noted as significant and p-values >0.05 are noted as highly significant.

On the other hand, three RCTs compared MIS-TLIF and open TLIF. Both open surgery and MIS were inspected for an average of 25 months (about 2 years). On postoperative day 2, blood loss and back pain were much lower in individuals receiving minimally invasive surgery (MIS), although it took significantly longer to supply intraoperative radiotherapy. Because all patients underwent open lumbar spine surgery prior to being randomly distributed to a group, the research had significant limitations. There was no difference in operating time, blood loss, or postoperative hospitalization between minimally invasive and open surgery. MIS needed less postoperative drainage and recovery time (40 days (about 1 and a half months) vs. 76 days (about 2 and a half months), but more time for intraoperative fluoroscopy. The VAS evaluations did not vary between the

two groups at 3, 6, 12, and 24 months (about 2 years) postoperatively, but the ODI did. Four years after surgery, a study of 41 patients (21 with MIS and 20 with open TLIF) found no clinically significant differences between the open TLIF and MIS groups. In terms of narcotic independence, two-year social cost, and return to work, numerous non-randomized prospective studies have set up that MIS-TLIF is better. MIS was comparable to open TLIF in terms of operating time, complication rate, and reoperation rate but with reduced blood loss and reduced hospital stay. In the second meta-analysis of 20+ trials, the surgical team received higher radiation. In a recent PLIF trial, it was found that MIS and open fusions have no differences. MIS had a higher revision and readmission rates than those in open TLIF and PLIF.

7. Conclusion

The strongest evidence does not confirm that the MIS process is better than open surgery for cervical or lumbar disc herniation. On the other hand, increased revision, readmission rates, reduced hospitalization, cost of treatment, and time to reach the normal life are disadvantages of MIS-TLIF fusion. No RCT comparing MIS-TLIF only to open PLIF has been conducted, which would supply valuable information. Any surgery involving MIS exposes the surgeon to much higher radiation. This is especially true for lumbar spine surgery, which is quite common. Whether this exposure had any effect on the patients is unknown at this time, and more research is needed. Patients must be informed of the most recent results to choose between MIS and open spine surgery or TLIF. This is critical given the current medical advertising climate that heavily promotes MIS.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] L. Lee, Z. Idris, T. Beng et al., "Outcomes of minimally invasive surgery compared to open posterior lumbar instrumentation and fusion," *Asian journal of neurosurgery*, vol. 12, no. 4, p. 620, 2017.

[2] A. J. Sayari, D. V. Patel, J. S. Yoo, and K. Singh, "Device solutions for a challenging spine surgery: minimally invasive transforaminal lumbar interbody fusion (MIS TLIF)," *Expert Review of Medical Devices*, vol. 16, no. 4, pp. 299–305, 2019.

[3] S. Arif, Z. Brady, Y. Enchev, N. Peev, and E. Encheva, "Minimising radiation exposure to the surgeon in minimally invasive spine surgeries: a systematic review of 15 studies," *Orthopaedics and Traumatology: Surgery & Research*, vol. 107, no. 7, Article ID 102795, 2021.

[4] G. De Biase, S. E. Gruenbaum, J. L. West et al., "Spinal versus general anesthesia for minimally invasive transforaminal lumbar interbody fusion: implications on operating room time, pain, and ambulation," *Neurosurgical Focus*, vol. 51, no. 6, p. E3, 2021.

[5] P.-I. Hung, M.-C. Chang, P.-H. Chou, H.-H. Lin, S.-T. Wang, and C.-L. Liu, "Is a drain tube necessary for minimally invasive lumbar spine fusion surgery?" *European Spine Journal*, vol. 26, no. 3, pp. 733–737, 2017.

[6] S. Sharif and A. Afsar, "Learning curve and minimally invasive spine surgery," *World neurosurgery*, vol. 119, pp. 472–478, 2018.

[7] F. C. Lovecchio, A. S. Vaishnav, M. E. Steinhaus et al., "Does interbody cage lordosis impact actual segmental lordosis achieved in minimally invasive lumbar spine fusion?" *Neurosurgical Focus*, vol. 49, no. 3, p. E17, 2020.

[8] H. G. David, N. D. Stekas, C. G. Varlotta et al., "Comparative analysis of two transforaminal lumbar interbody fusion techniques: open TLIF: versus: wiltse MIS TLIF," *Spine*, vol. 44, no. 9, pp. E555–E560, 2019.

[9] Y. Kotani, Y. Koike, A. Ikeura, H. Tokunaga, and T. Saito, "Clinical and radiologic comparison of anterior-posterior single-position lateral surgery versus MIS-TLIF for degenerative lumbar spondylolisthesis," *Journal of Orthopaedic Science*, vol. 26, no. 6, pp. 992–998, 2021.

[10] S. Ao, W. Zheng, J. Wu et al., "Comparison of preliminary clinical outcomes between percutaneous endoscopic and minimally invasive transforaminal lumbar interbody fusion for lumbar degenerative diseases in a tertiary hospital: is percutaneous endoscopic procedure superior to MIS-TLIF? A prospective cohort study," *International Journal of Surgery*, vol. 76, pp. 136–143, 2020.

[11] K. Chen, H. Chen, K. Zhang et al., "O-arm navigation combined with microscope-assisted MIS-TLIF in the treatment of lumbar degenerative disease," *Clinical Spine Surgery: A Spine Publication*, vol. 32, no. 5, pp. E235–E240, 2019.

[12] N. W. Jenkins, J. M. Parrish, E. D. K. Cha et al., "Validation of PROMIS physical function in MIS TLIF," *Spine*, vol. 45, no. 22, pp. E1516–E1522, 2020.

[13] B. Khechen, B. E. Haws, D. V. Patel et al., "Comparison of postoperative outcomes between primary MIS TLIF and MIS TLIF with revision decompression," *Spine*, vol. 44, no. 2, pp. 150–156, 2019.

[14] Y. Lv, M. Chen, S. L. Wang et al., "Endo-TLIF versus MIS-TLIF in 1-segment lumbar spondylolisthesis: a prospective randomised pilot study," *Clinical Neurology and Neurosurgery*, vol. 212, Article ID 107082, 2021.

[15] A. Hockley, D. Ge, D. Vasquez-Montes et al., "Minimally invasive versus open transforaminal lumbar interbody fusion surgery: an analysis of opioids, nonopioid analgesics, and perioperative characteristics," *Global Spine Journal*, vol. 9, no. 6, pp. 624–629, 2019.

[16] E. M. Soffin, D. S. Wetmore, J. D. Beckman et al., "Opioid-free anesthesia within an enhanced recovery after surgery pathway for minimally invasive lumbar spine surgery: a retrospective matched cohort study," *Neurosurgical Focus*, vol. 46, no. 4, p. E8, 2019.

[17] A. Minamide, M. Yoshida, A. K. Simpson et al., "Minimally invasive spinal decompression for degenerative lumbar spondylolisthesis and stenosis maintains stability and may avoid the need for fusion," *The Bone & Joint Journal*, vol. 100-B, no. 4, pp. 499–506, 2018.

[18] K. Mueller, D. Zhao, O. Johnson, F. A. Sandhu, and J.-M. Voyadzis, "The difference in surgical site infection rates between open and minimally invasive spine surgery for degenerative lumbar pathology: a retrospective single center

experience of 1442 cases," *Operative Neurosurgery*, vol. 16, no. 6, pp. 750–755, 2019.

[19] T. R. Deer, J. S. Grider, J. E. Pope et al., "The MIST guidelines: the Lumbar Spinal Stenosis Consensus Group guidelines for minimally invasive spine treatment," *Pain Practice*, vol. 19, no. 3, pp. 250–274, 2019.

[20] W. Zhao, C. Shen, R. Cai et al., "Minimally invasive surgery for resection of ossification of the ligamentum flavum in the thoracic spine," *Videosurgery and Other Miniinvasive Techniques*, vol. 1, no. 1, pp. 96–105, 2017.

[21] E. Darr, S. C. Meyer, P. Whang et al., "Long-term prospective outcomes after minimally invasive trans-iliac sacroiliac joint fusion using triangular titanium implants," *Medical Devices: Evidence and Research*, vol. Volume 11, pp. 113–121, 2018.

[22] T. R. Blattert, K. J. Schnake, O. Gonschorek et al., "Non-surgical and surgical management of osteoporotic vertebral body fractures: recommendations of the spine section of the German Society for Orthopaedics and Trauma (DGOU)," *Global Spine Journal*, vol. 8, no. 2_suppl, pp. 50S–55S, 2018.

[23] K. U. Lewandrowski, J. A. Soriano-Sánchez, X. Zhang et al., "Surgeon training and clinical implementation of spinal endoscopy in routine practice: results of a global survey," *Journal of spine surgery (Hong Kong)*, vol. 6, no. Suppl 1, p. S237, 2020.

Retraction

Retracted: Prevention of Business Risks of Internet Information Security Platforms Based on Blockchain Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] B. Yang, "Prevention of Business Risks of Internet Information Security Platforms Based on Blockchain Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7671810, 10 pages, 2022.

Hindawi

Research Article

Prevention of Business Risks of Internet Information Security Platforms Based on Blockchain Technology

Bingqing Yang (iD)

College of Information Engineering, Fuyang Normal University, Fuyang 236041, Anhui, China

Correspondence should be addressed to Bingqing Yang; 201207015@fynu.edu.cn

The increasing maturity of Internet information technology has led to the rapid rise of blockchain technology. In essence, a blockchain is a shared database. In recent years, blockchain technology has attracted the attention of the public and society. With the continuous development of the market economy, many domestic enterprises have grown in size. At the same time, companies are also faced with various risks, and business risks will have a great impact on the company's production. This article aims to prevent the business risk of Internet information security platform enterprises based on blockchain technology. Through the PoS algorithm, PoW algorithm, secure hash algorithm, and the principle of direct trust, the Internet information security platform is designed. Firstly, the platform is used for a business risk test with a small company, and the platform satisfaction survey is conducted on the employees of the company. The test results show that the subplatform design reduces the business risk of the company by 5%–10%, and employee satisfaction is high. The signature simulation of this algorithm under the same conditions as the SMRA and RSAR methods is also carried out. The results show that the performance of this algorithm is better. This study opens up a new path for small and medium-sized enterprises to prevent business risks.

1. Introduction

Due to the complexity of current market issues and fraud, some companies do not understand preventive measures and have little understanding of legal issues in their business, making them unable to effectively avoid operating risks. Blockchain is an important concept of bitcoin. In essence, it is a decentralized database [1]. In a narrow sense, blockchain is a chained data structure that combines data blocks in chronological order, and a distributed ledger that cannot be tampered with and forged in a cryptographic way [2]. Broadly speaking, blockchain technology is a new distributed infrastructure and computing method that uses blockchain data structures to verify and store data, uses distributed node consensus algorithms to generate and update data, uses cryptography to ensure the security of data transmission and access, and uses intelligent contracts composed of automatic script code to program and operate data.

The essence of blockchain technology is a decentralized database, and its main feature is to provide digital trust.

Information on the Internet is transparent and anonymous. It is this characteristic that makes information extremely insecure, and information loss, theft, and forgery are very common, which makes the Internet have some information security risks. If this problem cannot be resolved in time, it will seriously affect people's trust in the Internet and drive people away from the Internet. Only by comprehensively preventing the Internet information security platform from corporate management risks can we effectively guarantee Internet information security and reduce corporate operating risks. Improve the company's information risk management capabilities and raise the company's management level to the international advanced level, which provides a strong guarantee for the company's international development and cooperation. This paper aims to propose an effective business prevention scheme for the Internet platform based on blockchain technology to ensure the information security and transaction security of enterprises.

Technical personnel such as Viana E. proposed an Android-based solution for smartphones to identify and

report potentially unsafe settings and developed an application that can check smartphones to search for unsafe user configurations based on a predefined list. It reduces the attack of netizens by malicious users [3]. But at that time, they did not use the blockchain technology that has been particularly popular in recent years to reduce the risk of such Internet security information leakage. Researchers such as Bendul J C identified corporate risk culture and product vulnerability as the main factors influencing the implementation of preventive activities during transportation. Regression analysis illustrates the relationship between risk and quality-related influencing factors and the shipper's ability to perform risk prevention activities [4]. But they only studied the corporate risk culture and did not conduct in-depth research on the prevention of corporate operating risks. Scientists such as JJ Sikorski emphasized cryptocurrency mankind's research and practice potential in relation to the 4th industrial revolution. They came to the conclusion that bitcoin has the possibility to be underutilized in addition to supporting and boosting the industrial rebellion's efficiency, and they have defined the subject of future work [5]. However, they did not apply blockchain technology to Internet information security management and risk prevention in business operations. The above research contents are scientific and technological research on risk prevention, which can provide some references and ideas for this research. However, some research schemes do not use cutting-edge blockchain technology or do not protect the core content of enterprise risk. These are the key points of this research.

Protecting the security of information on the Internet, when using traditional information protection mechanisms, is difficult to ensure security. In order to better maintain the security of the network environment, the use of blockchain technology can enhance the protection effect, including user information, transaction activities, electronic communications, infrastructure, etc., which can be protected accordingly. A new blockchain-based Internet information security platform will improve information control capabilities and allow users to safely use online information. The experimental part of this paper not only pays attention to the maturity of technology but also has more humanistic characteristics, referring to the user's satisfaction with the platform.

2. Blockchain Security Algorithm

2.1. Direct Trust Mechanism.
The trust mechanism refers to the various parts of the company's system that constitute and influence the mutual trust relationship and the relationship management mechanism between them. Assuming that the lending decision of lender B is not affected by factors such as income and environment, but only depends on the credibility of the borrower, then, when the credibility is $P (0 \leq P \leq 1)$, the business operation risk rate is P.B. The basis for judging the credibility of the borrower's future performance of the contract is the borrower's credit information, where the information set displayed by the borrower W on the platform is $W = \{W_1, W_2, W_3, \ldots, W_i\}$. In actual

situations, W may deliberately conceal unfavorable information or exaggerate certain pieces of information, and its true information set is $W' = \{W'_1, W'_2, W'_3, \ldots, W'_i\}$. However, B can only see the set W, so the direct trust expectation made is

$$W = c_1 U_1 + c_2 U_2 + c_3 U_3 + \cdots + c_m U_n. \tag{1}$$

C_i $(i = 1, 2, \ldots)$ is the expected influence coefficient of each information factor on the credibility obtained by B through subjective judgment. The true value of direct trust is

$$W' = c'_1 U'_1 + c'_2 U'_2 + c'_3 U'_3 + \cdots + c'_m U'_n. \tag{2}$$

From the above two formulas, it can be concluded that the enterprise operating risk rate is

$$P(W > W' | j = k) = \frac{1 - \left(\sum_{j=0}^{k} \left(C_k^j \right)^2 / 4^j \right)}{2}. \tag{3}$$

2.2. PoW Algorithm.
PoW proves that a certain amount of work has been completed at the end of the work. The PoW algorithm has three main components: the work function proof, the retardation and difficulty values, and how to calculate the work function proof. And then use the PoW algorithm to block the input data. The most difficult part of PoW is the amount of calculation. The work end needs to do some difficult work to get a result, but the verifier can easily check whether the work end has done the corresponding work through the result. The benefits of the PoW algorithm include complete decentralization, free entry and exit of nodes, simple algorithms and simple implementations to reach consensus between nodes without exchanging other information [6, 7]. This algorithm is highly secure and requires a lot of damage to the system funds and allow 50% of the nodes in the entire network to fail. The PoW algorithm also has many shortcomings. For starters, it necessitates a significant amount of computational power, is susceptible to less supervision, and poses security hazards. Simultaneously, the introduction of huge mining tanks has brought attention to the issue of computational capacity, namely the potential of a "51 percent attack." The performance is low, but whenever a resolution is formed, the entire network is required to join in the computation. Reaching an agreement is impossible, disagreements are prone to occur, and a large number of confirmations must be waited for. The following are the specifications for the structure of chunk H said by a node in PoW: B is a particular hash technique, where T is a set quantity in $H(A)$target. That is, the hash value in its entirety. The chunk must be less than the stated amount and include a specific number of locations. The block is a lawful block such that if it fits this criterion, other nodes will accept it. If the node finds such a lawful zone, it will compensate the mines excavated with particular digital money. It also overcomes the problem of several distributed nodes causing problems with judgment. The overall chance P of discovering a valid chunk within every test, provided the greatest checksum is H, is

$$P = \frac{T_{\text{target}}}{H_{\text{max}}}. \tag{4}$$

Clinched alongside Bitcoin, the worth of t is balanced each 2016 obstructs (two weeks), which may be balanced toward the taking after formula

$$T_{(\text{new})} = \frac{T_{2016}}{2W} \cdot T_{\text{tar}}. \tag{5}$$

Around them, T_{2016} speaks to the duration of the time expended should produce the past 2016 territory squares. At those duration of the time expended may be shorter, those last focus quality will be Additionally more modest. Those challenge esteem of the created piece might make acquired toward recipe 3:

$$D_{(\text{difficulty})} = \frac{T_1}{T_{\text{current}}}. \tag{6}$$

2.3. PoS Algorithm. The PoW (proof of work) is the proof of workload, also known as mining. Most public chains or virtual currencies, such as bitcoin and Ethereum, are based on the PoW algorithm to realize their consensus mechanism. That is, the distribution of money is determined according to the effective work of the mining contribution. Compared with PoW, the PoS algorithm avoids a large amount of waste of resources caused by mining, reduces the consensus time between different nodes in a good low-level network environment, reaches the millisecond level, and has a large number of advantages. A PoS is a method of obtaining accounting rights based on decentralized computing competition. The supervisory ability is weak and has the same fault tolerance as PoW. It can be found in many different PoS applications that this is a hybrid way of using account balance custom mining (where b represents the account balance and t is a timestamp)

$$H(B, t) \le b * T_{(\text{target})}. \tag{7}$$

2.4. Blockchain Data Security Design. Blockchain is an important concept of bitcoin. In essence, it is a decentralized database. At the same time, as the underlying technology of bitcoin, it is a series of data blocks associated with cryptographic methods. Each data block contains a batch of information about bitcoin network transactions, used to verify the validity of its information (anticounterfeiting) and generate the next block. This article aims to improve the security and anonymity of blockchain transactions. For the signing algorithm of blockchain transactions, which cannot resist quantum computing attacks and the security problem of user identity privacy leakage, the research on key technologies for transaction security and privacy protection of blockchain has been carried out. The main research results of this paper are as follows:

Generation of key: randomly select two large prime numbers p and q to calculate: $n = pq$, $Q(n) = (p-1)(q-1)$, then select either positive integer e and calculate d so that

$de = 1 \bmod Q(n)$, where e is the public key and d is the private key.

The clear text encryption formula for the encryption process is

$$C = M^e \bmod n. \tag{8}$$

The ciphertext decryption formula for the decryption process is

$$M = C^d \bmod n. \tag{9}$$

The user uses the public key set L and his private key SK to generate a signature (r, s) as follows:

$$s = \left((1 + SK)^{-1} (k - rd) \right) \bmod n. \tag{10}$$

Using the sender node i message content and the public key as input to obtain the message verification code using formula (11), we get

$$\text{HMAC}(PK_i, S_i) = H(PK_i | S_i), \tag{11}$$

where PK is the public key of i, S is the message content, and H is the hash function. Blockchain users in this distributed network, the signature is verified in the following way, and the miner verifies the transaction through the following two formulas:

$$e \le 2s\sqrt{m}, \quad e \ne 0,$$
$$pk_a e = \sum_{i=1}^{d} (-1)^{M[i]} C_i. \tag{12}$$

The lattice signature algorithm used in transactions in the quantum blockchain scheme after this chapter satisfies the correctness requirement. The protocol process knows the tail signature $e = e' - u$. Thus, the following inequality can be obtained:

$$e \le e' + a \le 2s\sqrt{m}. \tag{13}$$

Since $A_1(e' - a) = (A_1 e' - A_1 a)$, we can output

$$e' \leftarrow \text{Sample Pre}(A_1, S_{ms}, b, s). \tag{14}$$

The signature satisfies $A_1 e' = b$. We get

$$\begin{aligned} A_1 e &= A_1(e' - a) \\ &= b - A_1 a \\ &= \sum_{i=1}^{d} (-1)^{M[i]} C_i + A_1 a. \end{aligned} \tag{15}$$

Through the analysis of equations (13) and (14), it is proved that the lattice signature algorithm used in the proposed postquantum blockchain scheme transaction satisfies the correctness.

ECC (error checking and correction) is error detection and correction algorithm for NAND. For the same key length, the key length of the ECC algorithm grows slowly, while the key length of the RSA algorithm grows exponentially, as shown in Table 1.

Table 1: Comparison of ECC and RSA key lengths.

Symmetric key length (bit)	RSA key length (bit)	ECC key length (bit)
86	1024	32
68	2048	16
16	3074	64
64	7660	4
54	14230	256

3. Design and Innovation of Internet Information Security Platform Based on Blockchain Technology

3.1. Blockchain Technology to Build a Distributed Model of Internet Information Security Platform. Blockchain technology creates a distributed model for specific applications. The distributed model places the data that different users need to access different parts on different servers to realize horizontal expansion. All users participate in the maintenance of the database. Use encryption technology to create data blocks and interconnect the data blocks [8, 9]. These interconnected data blocks are called blocks, and the data blocks are connected in chronological order to form a blockchain. Each block contains a time period, and the content in the block is the new information generated during the time period, and which covers the anticounterfeiting mark used to verify the information. During the information submission process, the information can only be retained in the database with the approval of the developer and cannot be deleted or changed at will. In terms of architectural design, the architecture of an Internet information security platform based on blockchain technology is shown in Figure 1.

3.2. Core Technology of Enterprise Operation Risk Prevention Based on Blockchain. The secure hash algorithm is an encryption technology [10]. It is to transform the input of any length into an output of a fixed length through a hash algorithm, and the output is the hash value. This transformation is a compression mapping; that is, the space of the hash value is usually much smaller than that of the input, and different inputs may be hashed into the same output. In short, it is a function that compresses messages of any length into a message digest of a fixed length. In the data calculation process, if the corresponding hash value is calculated in the forward direction, the calculation process is easy. When the hash value is used, the calculation of the corresponding data is an inverse calculation, and it can be known that this calculation is somewhat difficult. Blockchain networks use hash trees to perform calculations, which are used to construct a Merkle tree whose nodes are hash values. This method can verify the authenticity of a large amount of data. For example, in the security verification process, if a transaction is executed on the double flower subchain, the hacker will find a strong block in the network environment and create a weaker block with a higher cost before that. Judging from the displayed results, if the confirmation time

is 1 minute (in 1 MB blocks), a double-spending attack is required. Hackers may occupy 20% of the computing power of the entire network, and the hacking fee will reach 0.8 US dollars. If the block is 4 MB, the double-spending attack will need to pay a high transaction fee of more than $40. This type of security level is relatively low, and its security is similar to that of Bitcoin. Tables 2 and 3 show the transaction security of the security confirmation [11,12].

Analysis of the data in Tables 2 and 3 shows that as the information in the random block changes, its own Merkle hash also changes, and the information contained in the next block also changes. If hackers forge information, their computing power does not support false blocks, and the speed of false blocks exceeds the growth rate of the blockchain and is discarded [13, 14]. The information in each block is not only text information, but also data information, including nonlinear information (such as video information, image information, and various structured messages). The information stored in all blocks will be permanently protected and cannot be maliciously modified. Therefore, blockchain information is relatively safe [15–17].

4. Prevention of Business Risks of Internet Information Security Platform Enterprises

4.1. Necessity of Building an Internet Information Security Platform. Because this article uses the Fabric blockchain platform, it provides pluggable module configuration and management, by writing-related programs to implement the group signature scheme, and then using the RPC protocol in the blockchain network to remotely call the scheme of this article by controlling the upstream and downstream of the swarm nodes. We verify that our improved scheme can implement dynamic join and exit of nodes, as shown in Table 4.

After designing a blockchain-based Internet information security platform, we applied this platform to a small company to conduct a one-month business risk test. The test results are shown in Figure 2.

It can be seen from Figure 2 that in the first ten days of the year, the Internet information security platform was still in the debugging stage, and information security leaks were more serious. In particular, hackers attacked the system and employee login accounts were stolen. Vulnerability repairs have reduced the probability of various information security incidents by more than half in just over 10 days from early to late. This also proves that the Internet security platform can greatly reduce business risks, and the construction of an Internet information security platform is effective for the company. It is urgent to say [3, 18].

4.2. Corporate Employees' Satisfaction Survey on Internet Information Security Platforms. The data structure of blocks in a blockchain-based supply chain management system is shown in Figure 3.

The dotted box in Figure 3 corresponds to the data layer, the categorical Merkel tree layer, the classification Merkel
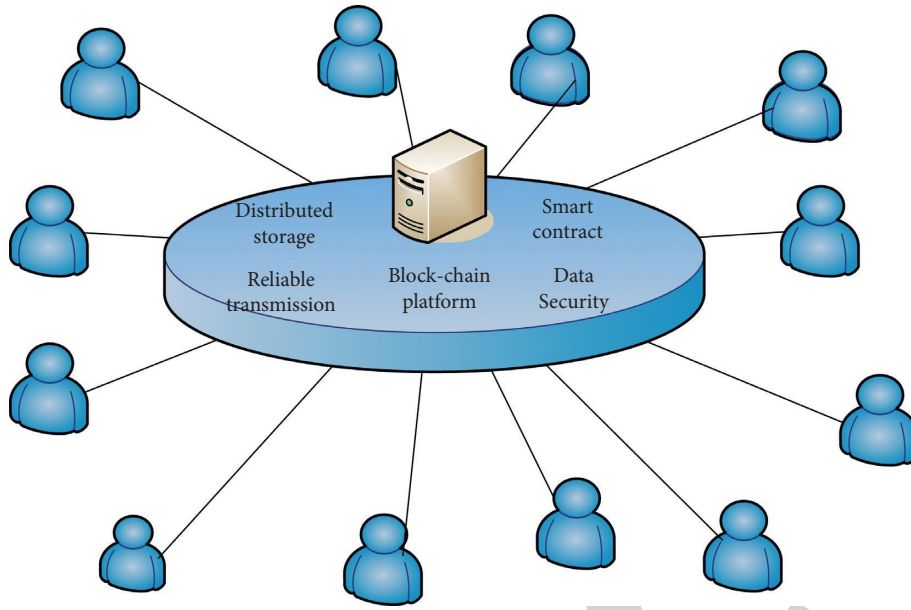
FIGURE 1: Internet information security platform based on blockchain technology.

TABLE 2: 5 seconds to confirm the security of the transaction after being confirmed in the double flower subchain.

The hash power of bad intruders is the mutual percentage in the network (%)	The estimated cost of a successful double-spending supply confirmed in 5 seconds (μB)		
	1.0 MB block	2.0 MB block	Probability of success (%)
0.05	600	1200	10
0.15	210	420	23.2
1.32	40	80	52.6
2.52	30	60	76.2
15	15	30	92.3
20	5	10	95.6

TABLE 3: 10 seconds to confirm the security of the transaction after being confirmed in the double flower subchain.

The hash power of bad intruders is the mutual percentage in the network (%)	The estimated cost of a successful double-spending supply confirmed in 10 seconds (μB)		
	2.5 MB block	5.0 MB block	Probability of success (%)
0.08	1500	3000	20
0.23	210	420	33.2
1.35	80	160	66.2
2.80	60	120	79.2
18.2	20	40	95.1
24.6	10	20	98.5

TABLE 4: Experimental environment configuration.

Configuration item	Configuration parameter
Operating system	Ubuntu 16.04 64 bit
CPU	Intel ® Xeon ® Platinum 8269CY 2.5 GHz
RAM	4G
Maximum network bandwidth	200 Mbps
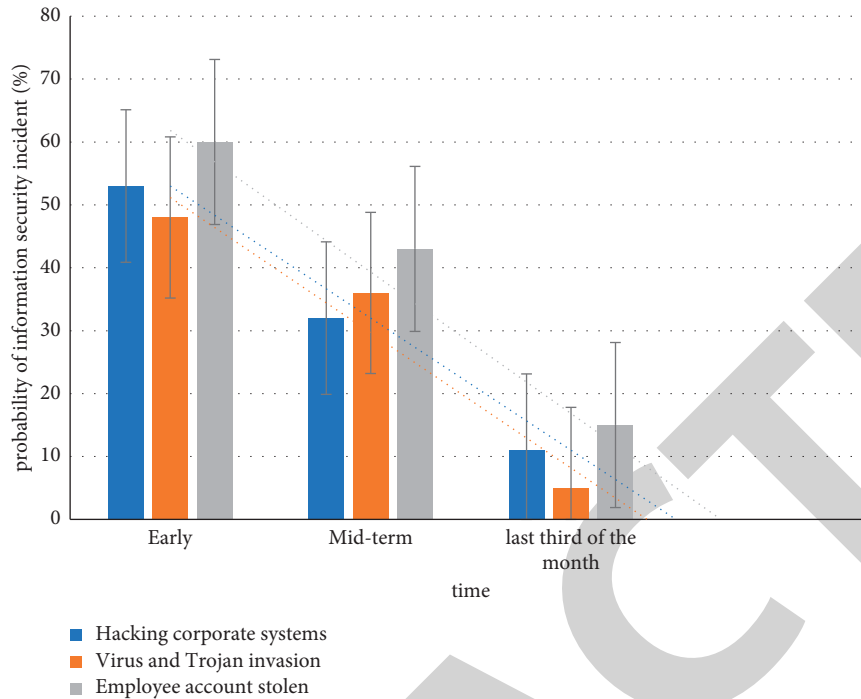
FIGURE 2: Business risk test.
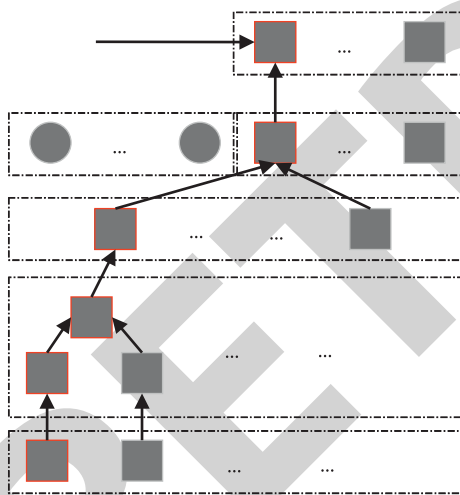
FIGURE 3: Data structure of blocks in a blockchain-based supply chain management system.

FIGURE 4: Blockchain-based virtual link network of information of a business entity in the supply chain.

tree root layer, the directory layer, and the block header layer in the block data structure, respectively.

For the supply chain with a production enterprise as the core, the operation process is described according to the order of business operation, so the process of generating the corresponding information for a certain business entity is also carried out according to the supply chain operation business process. The network of virtual links between a business principal's information in a blockchain-based supply chain is shown in Figure 4.

After a one-month risk test, we conducted a questionnaire survey on the company's employees of different age groups on their satisfaction with the platform design. We distributed 110 questionnaires and recovered 110, of which 2 were invalid questionnaires and the rest. There are 108 valid questionnaires in total, and the statistical results are shown in Figure 5.

It can be seen from Figure 5 that more than 50% of employees are very satisfied with the design of this security information platform. They believe that this platform can greatly reduce business risks and increase the company's net income. Less than 10% of employees are dissatisfied with the platform and think. This platform still has a lot of shortcomings, such as the design page is too simple, the system functions are not rich enough, etc.

4.3. Confidentiality Measures for Blockchain Electronic Signatures. For ring signatures, it has the following characteristics: *Correctness.* If the user can sign the signed message according to the correct signing steps and it has not been tampered with, then during the verification phase, the resulting ring signature must meet the ring signature verification formula. *Unconditional anonymity.* It is almost

Figure 5: Enterprise employee satisfaction survey.

impossible to determine the true identity of the signer. Assuming that there are n members participating in the ring signature, even if the attacker steals the private keys of all voters by some means, the chances of him being able to successfully identify the true signer of a signature will be less than $1/n$. *Unforgotten.* It is difficult to successfully forge a legitimate signature. In the whole process of ring signing, the verification process of each ring member is different, and the signature result is the same. In this case, it is difficult to counterfeit.

Blockchain technology is a decentralized distributed database technology with the characteristics of trustlessness, transparent transaction openness, and immutable data, which can effectively reduce data management costs, improve work efficiency, and protect data security. However, with the development of quantum computing, quantum computing attacks with strong computing power can crack classical cryptographic algorithms, which poses a huge threat to transaction security that relies on elliptic curve digital signature algorithms to ensure the transaction security of the blockchain. At the same time, due to the transparency of transaction information on the blockchain, relevant research has proved that there is also a risk of user identity privacy leakage [19]. Therefore, research on blockchain transaction security and privacy protection has become an important topic in the current blockchain security field.

Since the concept of electronic voting was first proposed in 1981, its basic model can be summarized as shown in Figure 6.

This paper designs an experiment to compare and analyze the signature generation times of the SMRA algorithm and the RSA-based signature algorithm RSAS, and the experimental data is shown in Table 5.
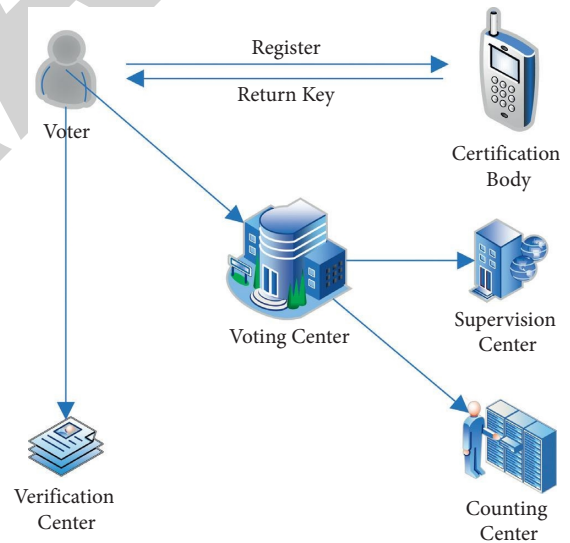
Figure 6: Diagram of the basic model of electronic voting.

The above data shows that in the case of an equal number of ring members, the signature generation time of the SMRA algorithm is less than that of the RSAS algorithm, and the signature efficiency is higher. Based on the above data, this article draws a simulation diagram as shown in Figure 7.

Through simulation experiments, the performance efficiency of the protocols is compared. The public key length, private key length, and signature length of the scheme in this section and the SMRA and RSAS methods are tested respectively, and $q = 2^{10}$, $d = 64$, and $m = 6n\log q$ are set under reasonable parameters according to the actual requirements in the lattice cryptographic algorithm, we set the security

TABLE 5: SMRA and RSAS signature time comparison analysis.

Number of ring members SMRA	Signature generation time (ms)	RSAS signature generation time (ms)
1	16.54	21.57
5	17.92	23.54
10	19.72	25.03
20	20.81	26.9
50	22.72	28.79
80	24.03	30.34
100	25.9	31.46
150	27.35	33.2
200	28.51	35.06
300	30.39	36.37
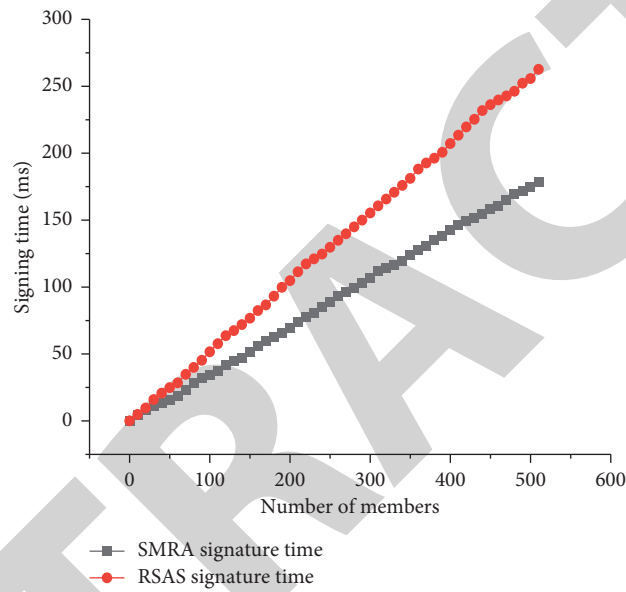
FIGURE 7: Simulation result graph.
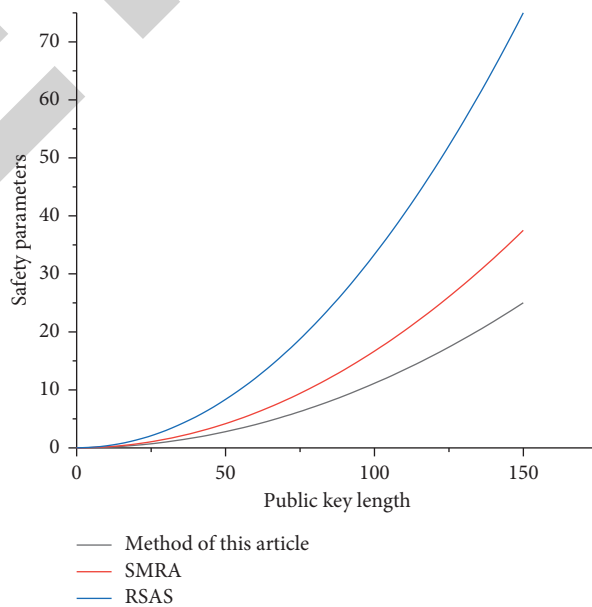
FIGURE 8: Public key length comparison result.
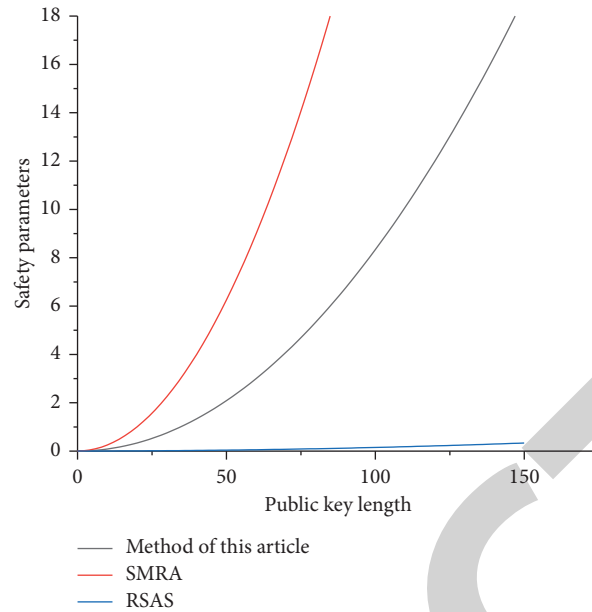
FIGURE 9: The result of the private key length comparison.
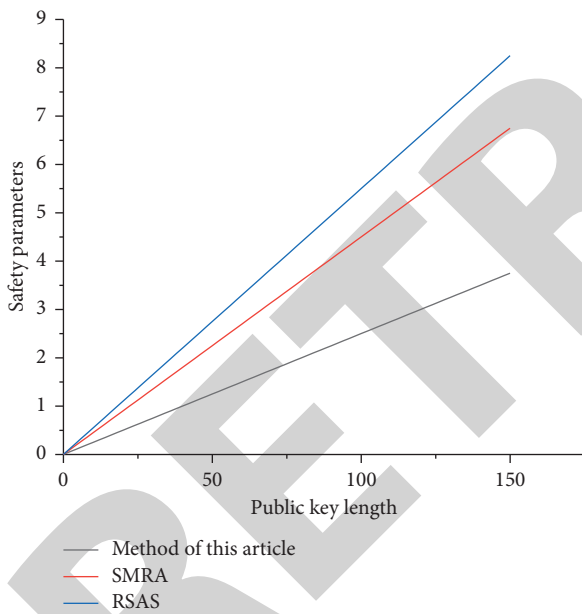
FIGURE 10: Signature length comparison result.

parameter n range from 10 to 160 increments, in turn, 16 tests. The experimental test results of public key length, private key length, and signature length for each scenario are shown in Figures 8–10, respectively.

Simulation test results show that with the continuous increase of the security parameter n, the public key, private key, and signature length of this section scheme are shorter than other schemes. In summary, the proposed signature algorithm has a relatively short one.

5. Conclusion

Internet information security issues are diverse, which means that Internet companies and users should have a sense of prevention when using the Internet and actively use new technologies to build an information security fortress on the Internet so as to reduce business risks. The development of blockchain technology provides a new model for Internet information security that is completely different from traditional forms of security prevention. The use of blockchain to develop Internet information security platforms can effectively protect the information security of enterprises. For example, a user's basic information, transaction information, and contact information can be guaranteed, thus greatly reducing the risks of business operations. But it should also be pointed out that any technology is a double-edged sword, with opportunities and challenges. The development of new technologies inevitably faces risks, and the blockchain is the same. Therefore, in the process of applying blockchain technology, we should think carefully, study carefully, make full use of the benefits of blockchain technology, and create an excellent new information security platform on the Internet that contributes to the prevention of business risks.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Hindawi

Research Article

Motion Fatigue State Detection Based on Neural Networks

Hu Li ⓘ, Yabo Wang, and Yao Nan

Lanzhou City College Sports Institute, Lanzhou 730070, China

Correspondence should be addressed to Hu Li; 201872397@yangtzeu.edu.cn

Aiming at the problem of fatigue state detection at the back of sports, a cascade deep learning detection system structure is designed, and a convolutional neural network fatigue state detection model based on multiscale pooling is proposed. Firstly, face detection is carried out by a deep learning model MTCNN to extract eye and mouth regions. Aiming at the problem of eye and mouth state representation and recognition, a multiscale pooling model (MSP) based on RESNET is proposed to train the eye and mouth state. In real-time detection, the state of the eye and mouth region is recognized through the trained convolution neural network model. Finally, the athlete's fatigue is determined based on PERCLOS and the proposed mouth opening and closing frequency (FOM). The experimental results show that in the training process, we set the batch_size = 100 and the initial learning rate = 0.01. When the evaluation index is no longer improved, the learning rate is reduced by 10 times to 0.001, and a total of 50 epochs are trained. The precision and recall of the system are high. Compared with the infrared image simulating the night state, the RGB image taken by the ordinary camera in the daytime has higher precision and recall. It is proven that the neural network has high detection accuracy, meets the real-time requirements, and has high robustness in complex environments.

1. Introduction

As shown in Figure 1, machine learning is the general name of a class of algorithms. It automatically analyzes a large amount of data to obtain new knowledge, accumulate experience, make it continuously improve its own performance, and use these laws to identify existing data for prediction or classification. In the fatigue motion detection algorithm based on image processing, machine learning algorithms are mainly used to detect and identify athletes' fatigue states [1]. In the process of fatigue motion detection, some classical machine learning algorithms are used. For example, the AdaBoost algorithm is used to detect the facial region, and then the eye and mouth images are obtained for subsequent processes. Support vector machine (SVM) algorithm is used to classify the manually extracted feature vectors so as to identify the eye state. Using neural network algorithms for face detection and keypoint location, constructing convolution neural network to automatically extract athletes' fatigue features, classify and recognize them, and so on. The other is to automatically extract the fatigue characteristics of athletes by using convolutional neural

networks and construct two convolutional neural network models to perform classification tasks for eye and mouth states, respectively, that is, open or closed states [2].

2. Literature Review

Liu Zhiqiang and others proposed a support vector machine (SVM) fatigue detection model based on the ASL eye tracker [3]. Wang Lei and others carried out a 36 h sleep deprivation experiment with an eye tracker to determine the thresholds of three fatigue judgment indicators: PERCLOS value, average eye closing time, and yawning frequency [4]. Nuevo and others used infrared devices to capture the eye movements of moving people and proposed a detection model based on AAM and PCA [5]. Jinjun Wang and others proposed a fatigue detection model based on the hidden Markov method based on simulated mover and infrared imaging [6]. Zeng Youwen and others also used contact instruments to collect physiological information from athletes. In the fatigue exercise experiment, they analyzed the correlation between EEG data and blink times and came to the conclusion that the EEG signal index and the blink time
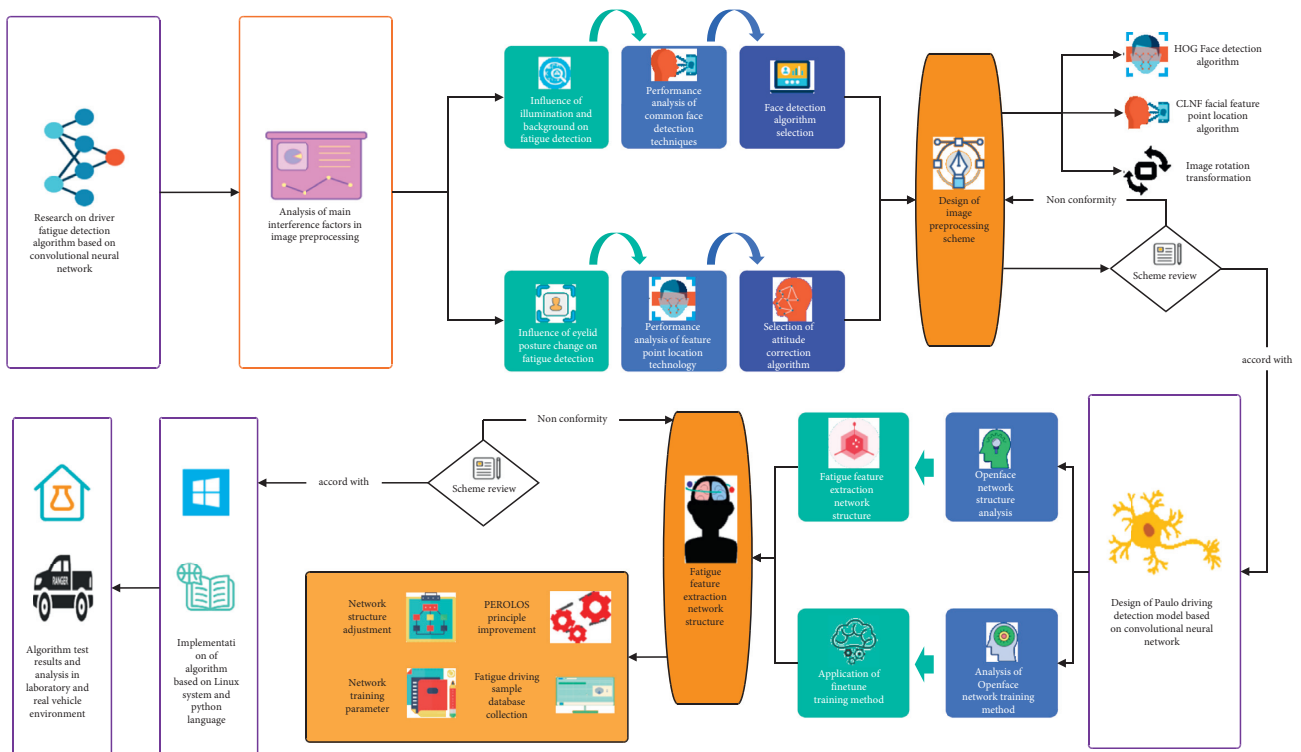
Figure 1: Technical roadmap of the paper.

index can maintain the same conclusion when indicating the fatigue degree of athletes. There are also many studies on how to use machine learning methods to analyze EEG signals [7]. Ye Chun and others designed a fatigue detection system for sportspeople based on Morlet wavelet theory and EEG signals, which has the advantages of fast operation speed and strong real-time performance [8]. Xie Zhi and others used the fusion K-means clustering method to process EEG signals. This method can fully consider individual differences and achieve an 80% recognition effect [9]. Kwok Tai Chui and others used the SVM algorithm to process EEG signals. The algorithm has less computation, fast running speed, and an average processing delay of only 0.55 ms [10]. Chin Teng Lin and others used the principal component analysis method to process EEG signals, which has a large amount of calculation and high accuracy [11]. Lee and others made innovations in physiological signal acquisition instruments, using watch-type instruments to collect information such as ECG signals and arm movement status of athletes for fatigue early warning. This acquisition equipment is smaller, easier to popularize, and can ensure that the routine operation of athletes is not affected [12]. Xintong and others used the AdaBoost algorithm for face detection, combined with the prior knowledge such as the geometric proportion of facial features to locate the coordinates of eyes and mouth, finally used the gray-scale integral projection method to extract the opening of eyes and the roundness of mouth as fatigue features, and determined fatigue according to the PERCLOS principle [13]. Wu Minjie and others adopted a similar algorithm design idea, combined with the AdaBoost algorithm and template matching algorithm to

complete the function of fatigue feature extraction, and introduced the frequency of mouth opening and closing as one of the criteria into the fatigue detection system of sportspeople [14]. You Feng and others, as shown in Figure 2, after using the AdaBoost algorithm to detect a face, they use an elliptic curve fitting method to obtain eye-opening so as to judge fatigue [15].

Fatigue detection based on visual image processing has been a research topic for many years. However, due to the large interference of ambient light and human expression in the real scene, the accuracy of face detection and facial behavior feature detection still needs to be improved. To solve this problem, this paper proposes a cascade deep learning structure and a fatigue state detection model based on multiscale pooling networks (MSP net) [16]. The trained MTCNN model is used to detect the face, extract the data of eyes and mouth, put the data of eyes and mouth into the trained MSP net CNN for detection, judge the opening and closing of eyes and mouth, and finally jointly judge fatigue according to PERCLOS and frequency of open mouth (FOM). The specific process is shown in Figure 3.

3. Neural Network MSP Net

3.1. Pretreatment

3.1.1. Data Acquisition. The fatigue detection system in this paper adopts the cascade form of two deep learning models; one is MTCNN for detecting faces and extracting eye and mouth images, and the other is a multiscale pooled CNN for judging the state of eye and mouth images as proposed in
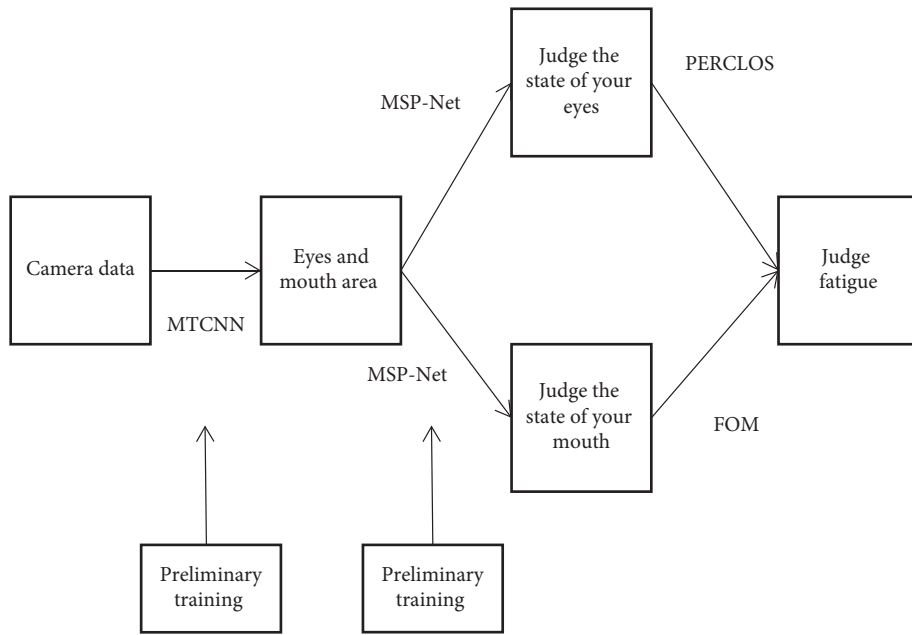
Figure 2: AdaBoost algorithm detection.
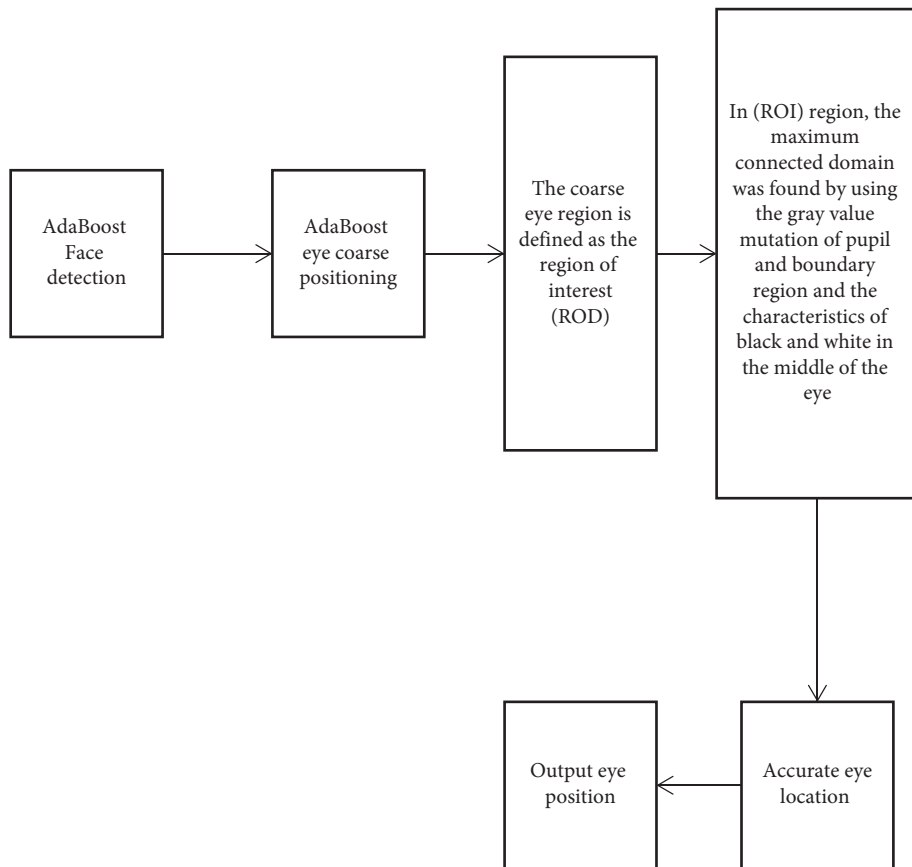
Figure 3: Algorithm flow.

this paper. Both networks require pretraining. Because the deep learning model has made remarkable achievements in face detection in recent years, this paper uses the existing excellent network to obtain the face and eye mouth area and focuses on the fatigue detection task [17]. The MTCNN model used in this paper is a trained model. The author uses the widerface database. There are about 200000 face images, including frame annotation and five key point information

(center coordinates of two eyeballs, nose tip coordinates, and two mouth corner coordinates). The regression of network training is the face frame and the coordinates of five key points. The data of eye and mouth image state judgment network training are independently collected by the author in practical work, including some actual athlete images and 21 volunteers. Considering that athletes have night training, the acquisition equipment uses infrared cameras in addition to ordinary cameras. In the process of image and video acquisition, various complex environmental problems of athletes during actual sports are comprehensively considered. The collected image data includes eyes closed, mouth open and closed, no glasses, wearing glasses, front and side, etc. After collection, we filter and eliminate some noise samples. Finally, the selected samples are classified as the opening, closing, opening and closing of ordinary cameras and the opening, closing, opening and closing of infrared cameras. Among them, there are 36764 samples in the eyes and 15185 samples in the mouth [18].

3.1.2. Data Preprocessing.
Because the brightness of the sample image collected by the infrared camera is generally low, the effect of directly putting it into the network training is not good, so the sample needs to be preprocessed. This paper uses the histogram equalization method. The purpose of histogram equalization is to enhance the local contrast. Its main steps are to calculate the cumulative probability of the gray level of the original image and map the original gray value to the new gray value according to the mapping relationship. For example, the gray-scale mapping relationship of gray-scale images with gray-scale values of 0–255 is as follows:

$$p_k = 255 \times \sum_{i=0}^{k} \frac{n_i}{n}, \quad k = 0, 1, 2, \ldots, L-1, \tag{1}$$

where p_k represents the gray value after mapping; k represents the kth gray level; L gray levels in total; n represents the number of all pixels; n_i represents the number of pixels of the i-th gray level; and the resulting p_k is rounded at the end.

3.2. Face Detection.
Traditional face detection algorithms, such as AdaBoost or frame difference, have poor robustness to complex environments. A deep learning model has great advantages in this regard. In this paper, the MTCNN model is used for face detection. The network structure of MTCNN is mainly divided into the following three layers:

(a) *P-Net*. Firstly, the image pyramid is constructed, and then the candidate face windows and bounding box reregression vectors are obtained through a fully convolutional network (FCN), which are used to calibrate the candidate face windows. Then, non-maximum suppression (NMS) is used to merge highly coincident candidate regions [19].

(b) *R-Net*. We put the candidate areas obtained by p-net into this network to further screen out a large number of wrong candidate areas and perform

calibration. Finally, NMS is also used to merge candidate areas.

(c) *O-Net*. This layer of the network is similar to R-Net, but this layer is made more detailed, and the candidate areas will be more strictly supervised. Finally, five more key point coordinates will be output [19].

MTCNN has good robustness and can still accurately detect the face rotated at a certain angle. The face is detected by MTCNN, and the eyes and mouth can be successfully marked according to the returned five key points.

3.3. Model Design.
The multiscale pooled convolutional neural network (MSP net) model is improved on the basis of the structure of RESNET, retains the concept of residual, and modifies the original max pooling to the multiscale pooling (MSP) proposed in this paper at the pooling layer to improve the recognition effect of images collected at different resolutions [20]. The structure of MSP is shown in Figure 4.

The steps of the MSP module are as follows: (a) first passed through twice max pooling to obtain a group of feature maps with side lengths four times smaller than the original feature map. (b) The original feature map is scaled, the side length is doubled, and then the new feature map is max pooled to obtain another group of feature maps with a side length of 1/4 of the original feature map. (c) Scaled the original feature map again, and the side length is reduced to 1/4 times of the original. (d) The three output feature maps are cascaded and introduced into the subsequent deep learning network [21]. The idea of a multiscale pooling model comes from the spatial pyramid pooling model. Compared with the spatial pyramid model, the advantage of multiscale pooling is that its substitution position is more flexible and can be used many times at the beginning, middle, or end of the network.

The structure of the MSP net network designed in this paper is shown in Figure 5. The training methods for the eyes and mouth are the same. Here, we take eyes as an example. In MSP net, the input image is 48×48 sizes gray image, after one convolution and MSP output $12 \times 12 \times 48$. The convolution kernel size is 3×3. After that, a residual block is passed. There are two layers of convolution in the residual block, and the size of the convolution kernel is still 3×3. Residual block output is $12 \times 12 \times 48$ feature maps. After another −max pooling, the output is $6 \times 6 \times 48$ feature maps. Then, the feature maps are converted into a one-dimensional vector and entered into the full connection layer. The input sequence length of the full connection layer is 1728, and there is a hidden layer with a length of 1000. Finally, the classification results are output by Softmax. The categories are divided into four categories: ordinary camera opening, infrared camera opening, ordinary camera closing, and infrared camera closing.

In order to verify the effectiveness of the MSP net, this paper not only uses the MSP net for experiments but also uses the classical Alex net and RESNET structures as comparison networks for training, testing, and related comparison experiments.

Characteristics of figure

FIGURE 4: Multiscale pooling (MSP) module.

FIGURE 5: MSP net structure.

3.5. Fatigue State Detection

3.5.1. PERCLOS.
PERCLOS is the ratio between the number of eye closure frames per unit time and the total number of frames per unit time. The calculation formula is as follows:

$$f_{\text{per}} = \frac{n}{N} \times 100, \tag{5}$$

where n represents the number of closed eye frames and N represents the total number of frames per unit time. PERCLOS can well quantify the degree of athletes' eye closure. When PERCLOS reaches a certain threshold, it can be judged that the athletes' eye closure time is too long, and it can be preliminarily considered that they have entered a fatigued state.

3.5.2. Mouth Opening and Closing Frequency.
FOM, similar to PERCLOS, represents the ratio between the number of frames with mouth open per unit time and the total number of frames per unit time. The calculation formula is as follows:

$$f_{\text{fom}} = \frac{n}{N} \times 100. \tag{6}$$

Like PERCLOS, n represents the number of shut-up frames and N represents the total number of frames per unit time. The greater the value of these two indicators, the greater the degree of fatigue. The final fatigue state detection needs to be considered together.

3.5.3. Fatigue State Detection.
After the pretraining of all deep networks is completed, the threshold of fatigue state is set according to PERCLOS and FOM, and the whole network system is applied to real-time detection. The specific steps are as follows: the camera captures the athlete's video, and MTCNN captures the face and five key points in each frame and extracts the eye and mouth areas. An MSP net is used to detect the state of the eyes and mouth captured in each frame and save them in a fixed long queue. The algorithm detects the change of the median value in the queue in real-time. When the distribution of all values in the queue reaches the threshold fatigue state, the alarm mechanism starts to remind athletes that they have entered fatigue.

3.4. Loss Function and Optimization Method.
Finally, the network uses the Softmax classification method, which is divided into four categories. The SoftMax function is defined as follows:

$$p_j = \frac{\exp\left(y'_j\right)}{\sum_{k=0}^{3} \exp\left(y'_k\right)}, \quad j = 0, 1, 2, 3, \tag{2}$$

where p_j represents the probability of class j; $y'_j = \sum_i h_i w_{i,j} + b_j$ represents the output of the last layer of the full connection layer, h_i is the output of the previous layer, and $w_{i,j}$ and b_j are the weight and offset of the last layer, respectively. The loss function is defined as cross-line and expressed as

$$L_m = -\sum_{j=0}^{3} 1\{y = j\}\log p_j, \tag{3}$$

where L_m represents the cross-line of the m-th sample; $1\{y = j\}$ represents the indicative function, that is, when $y = j$, the function is 1, and when $y \neq j$, the function is 0. (3) is the loss function of a single sample. When there are M training samples, the loss function needs to be averaged and it is expressed as

$$L = \frac{1}{M} \sum_{m=1}^{M} L_m. \tag{4}$$

The optimization method uses adaptive motion estimation (Adam).

FIGURE 6: Model loss on eye dataset model_loss.
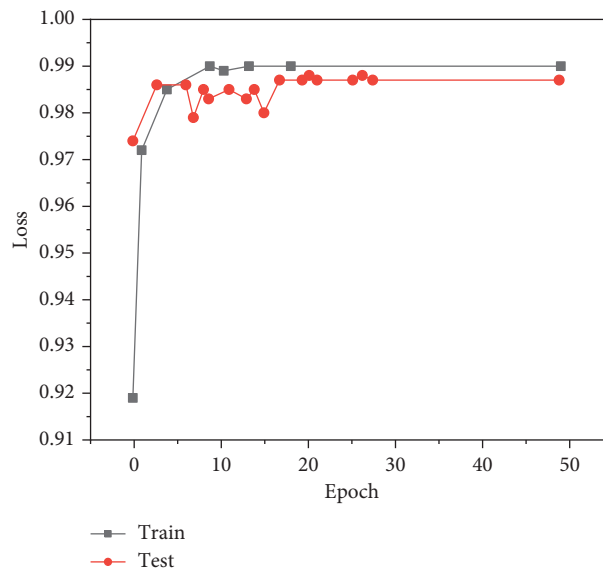
FIGURE 7: Model accuracy on eye dataset model_accuracy.

4. Experimental Results and Analysis

This paper constructs a neural network model based on the deep learning framework and uses the framework of multithreaded input data provided by tensor flow to combine the training data (batch) and disrupt the data order, which can improve the efficiency of model training. During training, we set the batch_ size = 100 and learn_ rate = 0.01. When the evaluation index is no longer improved, the learning rate decreases 10 times to 0.001 and the lowest to 0.00001. One epoch means that all training sets are trained once. In this paper, 50 epochs are trained on the training set, the Adam adaptive moment estimation optimization algorithm is used for backpropagation, the callback function is set to monitor the loss val_loss of the verification set, and the network model with the minimum loss on the verification set is saved as the final training model. The graphs of training loss and verification loss val_loss are shown in Figure 6, and the graphs of accuracy rate train accuracy on training set and accuracy rate val_accuracy on verification set are shown in Figure 7.

The verification dataset and test dataset are used to test the model saved after training. The detection accuracy on each dataset is shown in Table 1.

In order to verify the performance of the algorithm proposed in this paper, using the public CEW dataset, our algorithm is compared with the eye state recognition algorithm proposed in recent years. The detection accuracy is shown in Figure 8.

TABLE 1: Detection accuracy on each eye dataset.

Dataset	CSW data	Yaw DD data (%)	Self-built data (%)	Average detection accuracy (%)
Validation set	0	98.7	98.6	98.65
Test set	97.87%	98.2	98.44	98.17
Average detection accuracy	97.87%	98.45	98.54	98.42
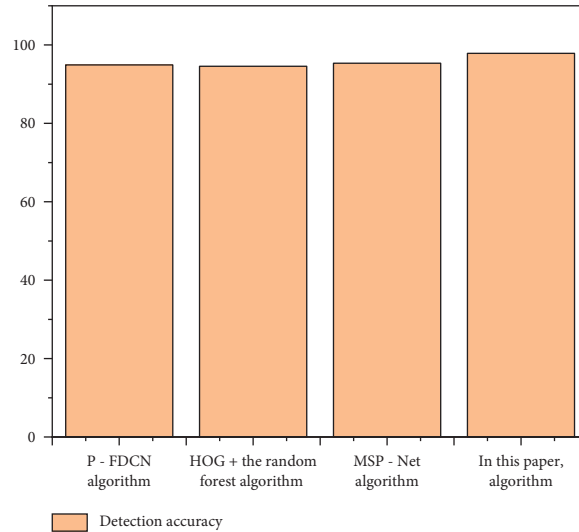
FIGURE 8: Performance comparison of different eye state detection algorithms.
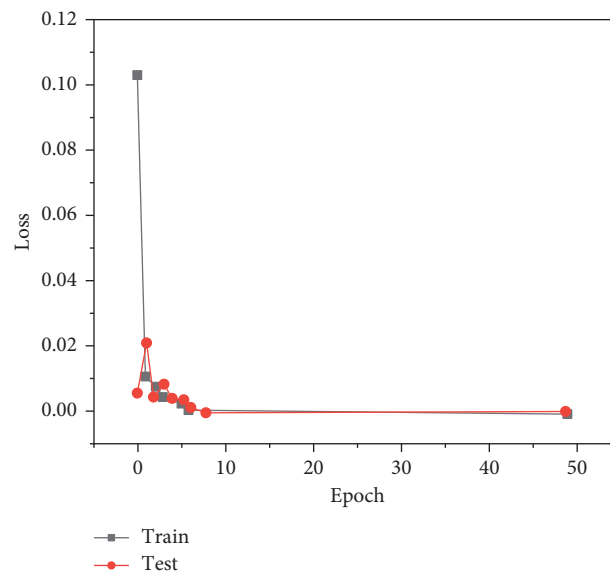
FIGURE 9: Model loss on dataset model_loss.

The algorithms proposed in the above other references collect monocular images to analyze the eye state, while our algorithm analyzes the athlete's eye state through binocular images. Compared with monocular images, the proposed algorithm can extract more abundant features and has higher detection accuracy. The images in the CEW dataset are ideal, and there are few postures such as head deflection and tilt. The experimental results show that our algorithm has high detection accuracy compared with other algorithms. In the yaw

DD dataset, the head posture of athletes is changeable, and there will be inclination, deflection, and other situations, where the detection effect of other algorithms will be worse.

After preparing the training dataset, the next step is to train the network and verify the performance of the network model. This paper uses the framework of multithreaded input data provided by tensor flow to combine training data (batch), which also disrupts the order of training data. In the training process, we set the
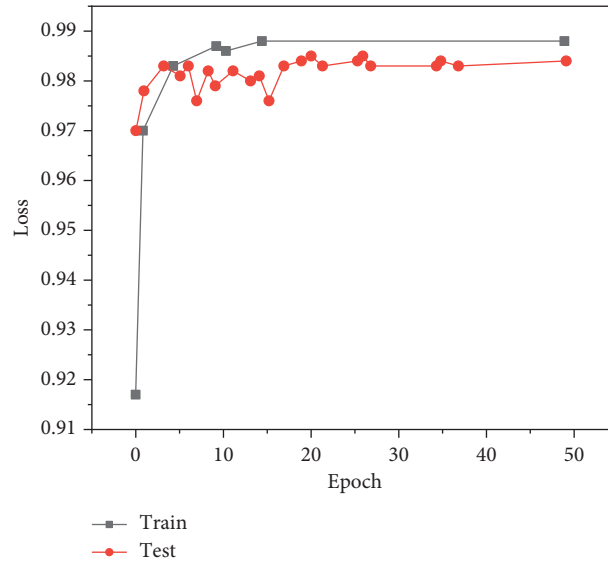
FIGURE 10: Model accuracy on mouth dataset model_accuracy.
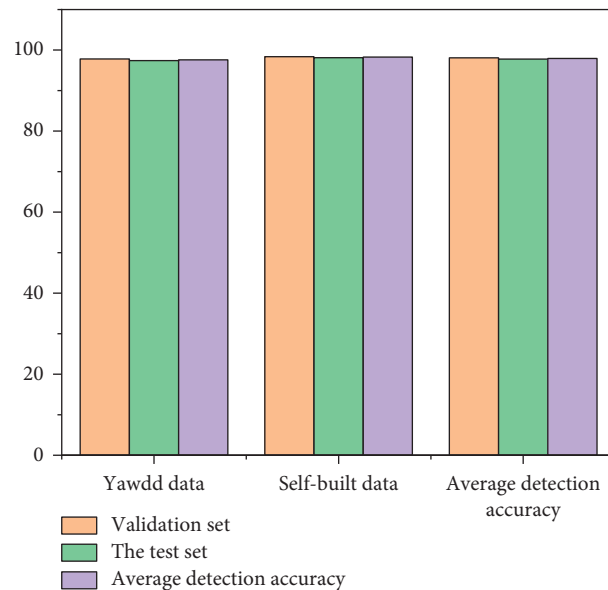
FIGURE 11: Detection accuracy on each mouth dataset.

batch_size = 100 and learn_rate = 0.01. When the evaluation index is no longer improved, the learning rate decreases 10 times to 0.001, and a total of 50 epochs are trained. In this paper, the Adam adaptive optimization algorithm is used for backpropagation, the callback function is used to monitor the loss of the verification set, and the network model with the least loss on the verification set is saved as the final training model. The graphs of training loss train_loss and verification loss val_loss are shown in Figure 9, and the graphs of accuracy on the training set and accuracy val_accuracy on verification set are shown in Figure 10.

The model saved after training is tested with the mouth verification dataset and the test dataset. The detection accuracy on each data set is shown in Figure 11.

In order to verify the performance of the algorithm proposed in this paper, using the public yaw DD dataset, our algorithm is compared with the mouth state recognition algorithm proposed in recent years. The experimental results are shown in Figure 12. Compared with other mouth state recognition algorithms, the algorithm proposed in this paper performs mouth correction, so the detection accuracy is higher when the athlete's head posture is tilted and deflected. Experimental results also show that the proposed algorithm is better than other algorithms.

It can be seen from Figure 13 that videos of athletes under normal exercise, talking/laughing, and yawning are selected from the test dataset. The following two figures are the results of analyzing the fatigue degree of athletes through the state of their mouth. When the system detects that the
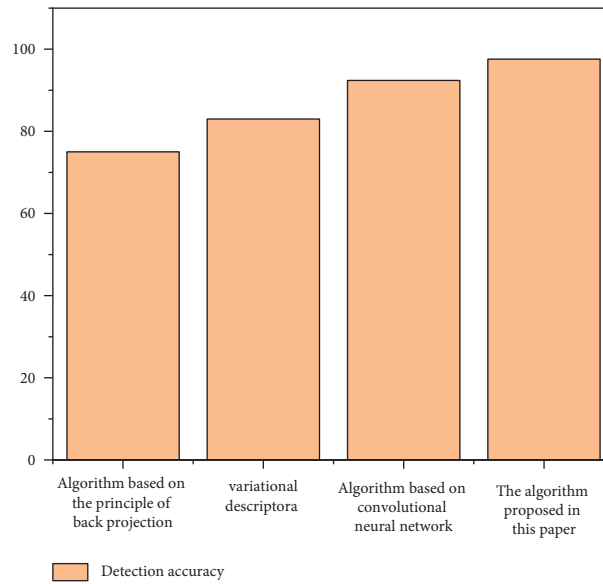
FIGURE 12: Performance comparison of different mouth state detection algorithms.
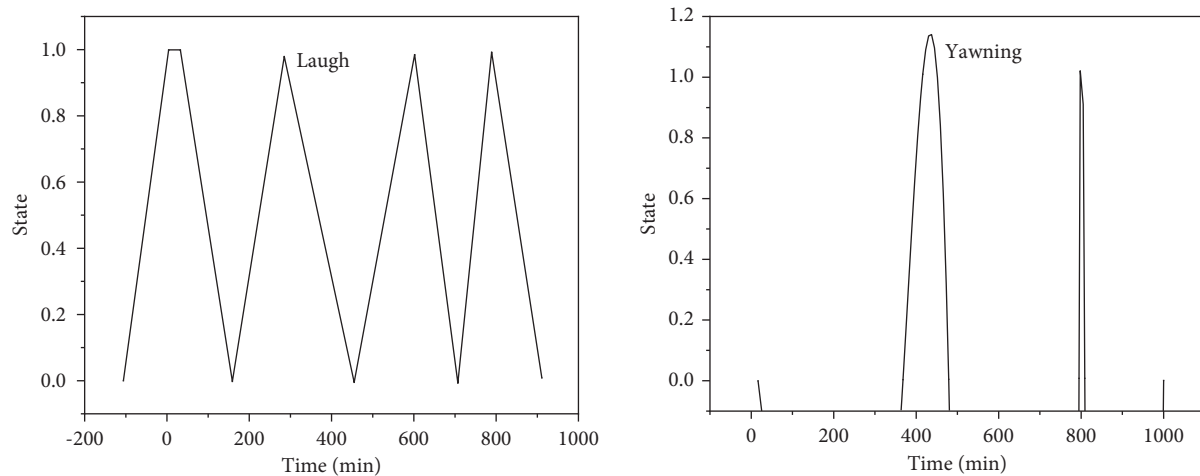
FIGURE 13: Two different mouth states: (a) Normal and (b) Yawning.

athlete is yawning (Figure 13), an audible warning is issued to remind the athlete to pay attention.

5. Conclusion

In this paper, a cascade deep learning structure and a real-time fatigue detection system based on multiscale pooled convolutional neural networks are designed. Firstly, the athlete's face is detected by MTCNN to extract the key positions of eyes and mouth. Then the eye and mouth images are sent to the multiscale pooled MSP net for the state test, and a fixed-length queue is set. The queue saves the detection results of each frame in unit time, and the fatigue state is judged jointly by PERCLOS and mouth opening and closing frequency (FOM). Experiments show that the proposed algorithm has high detection accuracy, can achieve the effect of real-time detection, and has high robustness in complex environments. The method proposed in this paper will be further transplanted and optimized for the embedded platform.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Luo, T. Qiu, C. Liu, and P. Huang, "Research on fatigue driving detection using forehead eeg based on adaptive multiscale entropy," *Biomedical Signal Processing and Control*, vol. 51, pp. 50–58, 2019.

[2] Y. Wang, R. Huang, and L. Guo, "Eye gaze pattern analysis for fatigue detection based on gp-bcnn with esm," *Pattern Recognition Letters*, vol. 123, pp. 61–74, 2019.

[3] Z. Cui, H. M. Sun, R. N. Yin, L. Gao, H. B. Sun, and R. S. Jia, "Real-time detection method of driver fatigue state based on deep learning of face video," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 25495–25515, 2021.

[4] C. Song, P. Zeng, Z. Wang, T. Li, L. Qiao, and L. Shen, "Image forgery detection based on motion blur estimated using convolutional neural network," *IEEE Sensors Journal*, vol. 19, no. 23, pp. 11601–11611, 2019.

[5] X. Zeng, Y. Wu, Z. Yang, Y. Shen, and L. Ma, "Research on high-frequency small-load fatigue testing device driven by piezoelectric actuator," *Review of Scientific Instruments*, vol. 90, no. 7, Article ID 076102, 2019.

[6] E. R. Monteiro, P. B. Costa, V. G. Corrêa Neto, B. J. Hoogenboom, J. Steele, and J. d. Silva Novaes, "Posterior thigh foam rolling increases knee extension fatigue and passive shoulder range-of-motion," *The Journal of Strength & Conditioning Research*, vol. 33, no. 4, pp. 987–994, 2019.

[7] X. Hu and G. Lodewijks, "Exploration of the effects of task-related fatigue on eye-motion features and its value in improving driver fatigue-related technology," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 80, no. 3–4, pp. 150–171, 2021.

[8] X. Chai, Z. Zhang, K. Guan, T. Zhang, J. Xu, and H. Niu, "Effects of fatigue on steady state motion visual evoked potentials: optimised stimulus parameters for a zoom motion-based brain-computer interface," *Computer Methods and Programs in Biomedicine*, vol. 196, no. 6, Article ID 105650, 2020.

[9] N. Tang, "Image recognition algorithm for exercise fatigue based on fpga processor and motion image capture," *Microprocessors and Microsystems*, vol. 81, no. 3, Article ID 103756, 2020.

[10] E. Franchini, M. Y. Takito, E. D. Alves, S. A. Shiroma, U. F. Julio, and C. Humberstone, "Effects of different fatigue levels on physiological responses and pacing in judo matches," *The Journal of Strength & Conditioning Research*, vol. 33, no. 3, pp. 783–792, 2019.

[11] Z.-K. Gao, Y.-L. Li, Y.-X. Yang, and C. Ma, "A recurrence network-based convolutional neural network for fatigue driving detection from eeg," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 11, Article ID 113126, 2019.

[12] M. D. Basar, A. D. Duru, and A. Akan, "Emotional state detection based on common spatial patterns of eeg," *Signal, Image and Video Processing*, vol. 14, no. 3, pp. 473–481, 2020.

[13] Z.-f. Zhang, Y.-l. Liang, T.-y. Lü et al., "Effect of moxibustion at shenque (cv 8) on myocardial remodeling and function in exercise-induced fatigue rats," *Journal of Acupuncture and Tuina Science*, vol. 19, no. 4, pp. 249–257, 2021.

[14] Q. Guo, X. Guo, J. Fan, and C. Wu, "Research on high-cycle fatigue behavior of fv520b steel based on intrinsic dissipation," *Acta Metallurgica Sinica*, vol. 90, no. 4, pp. 248–255, 2015.

[15] W. H. Gu, Y. Zhu, X. D. Chen, L. F. He, and B. B. Zheng, "Hierarchical CNN-based real-time fatigue detection system by visual-based technologies using MSP model," *IET Image Processing*, vol. 12, no. 12, pp. 2319–2329, 2018.

[16] Y. Yuan, Z. Jiang, H. Zhang, D. Zhao, and B. Cai, "Ship detection in optical remote sensing images based on deep convolutional neural networks," *Journal of Applied Remote Sensing*, vol. 11, no. 4, p. 1, 2017.

[17] X. Zhou and P. Jiang, "Variation source identification for deep hole boring process of cutting-hard workpiece based on multi-source information fusion using evidence theory," *Journal of Intelligent Manufacturing*, vol. 28, no. 2, pp. 255–270, 2017.

[18] W. Yan, G. Xu, M. Li et al., "Steady-state motion visual evoked potential (ssmvep) based on equal luminance colored enhancement," *PLoS One*, vol. 12, no. 1, Article ID e0169642, 2017.

[19] C. Han, G. Xu, J. Xie, C. Chen, and S. Zhang, "Highly interactive brain-computer interface based on flicker-free steady-state motion visual evoked potential," *Scientific Reports*, vol. 8, no. 1, p. 5835, 2018.

[20] J. Ge and H. Gu, "Cable motion capture and analysis based on optical tracking system," *Robotica*, vol. 33, no. 4, pp. 1–12, 2015.

[21] J. Xie, G. Xu, J. Wang, M. Li, C. Han, and Y. Jia, "Effects of mental load and fatigue on steady-state evoked potential based brain computer interface tasks: a comparison of periodic flickering and motion-reversal based visual attention," *PLoS One*, vol. 11, no. 9, Article ID e0163426, 2016.

Research Article

Big Data Analysis and Prediction System Based on Improved Convolutional Neural Network

Xuegong Du ⬡, Xiaojun Cao, and Rui Zhang

College of Information Engineering, Lanzhou University of Finance and Economics, Lanzhou 730020, China

Correspondence should be addressed to Xuegong Du; duxg@lzufe.edu.cn

This paper presents a big data analysis and prediction system based on convolutional neural networks. Continuous template matching technology is used to analyze the distributed data structure of big data, and the information fusion processing of cloud service combination big data is combined with matching related detection methods, frequent item detection, and association rule feature extraction of high-dimensional fusion data. A clustering method is adopted to realize the classification and mining of cloud service portfolio big data. The hardware equipment of the car to detect the surrounding environment is complicated, and the combination of the convolutional neural network and the camera to detect the surrounding environment has become a research hotspot. However, simply using the convolutional neural network to process the camera data to control the turning angle of the car has the problems of long training time and low accuracy. An improved convolutional neural network is proposed. The experimental results show that the accuracy of data mining by this method is 12.43% and 21.76% higher than that of traditional methods, and the number of iteration steps is shorter, indicating that the timeliness of mining is higher. This network structure can effectively improve the training speed of the network and improve the accuracy of the network. It is proven that the convolutional neural network has faster training speed and higher accuracy.

1. Introduction

All walks of life have begun to carry out artificial intelligence research in an all-around way, the most critical of which is the deep learning technology (Figure 1). "Deep learning" is a multilayer neural network, and "deep" in a sense refers to the number of layers of artificial neural networks. This is a brand-new field in machine learning research. This method aims to simulate human's intelligent behavior by simulating human thinking process so that, after training, the machine can show intelligent behavior that looks like a human, so that the ability of machine learning can be displayed. There is a possibility of surpassing human intelligence [1]. Deep learning has led to the rise of artificial intelligence. Traditional artificial intelligence algorithms rely on artificially summarized rules to program solutions to problems. But deep learning is different. It does not require artificial extraction of the characteristics or rules of the problem. It can learn from the input. A large amount of data spontaneously summarizes the law, adaptively adjusts its own structure so as to draw inferences from one another and generalize it to a case that has never been seen before [2]. To sum it up in one sentence, the most important feature of deep learning is that it can automatically learn W from data. Basic deep learning models can be divided into two categories: generative models and discriminative models. The former mainly includes restricted Boltzmann machine (RBN) models, autoencoder (AE) models, and deep belief network (DBN) models, which are generally used to express high levels of data. Order correlation or joint statistical distribution describing data; the latter mainly include convolutional neural network (CNN) model, recurrent neural network (RNN) model, deep stacking network (DSN) models, and long short-term memory network models, are usually used to classify the internal pattern of the data or describe the posterior distribution of the data [3].

The convolutional neural network is the most widely used deep learning network in computer vision. It has
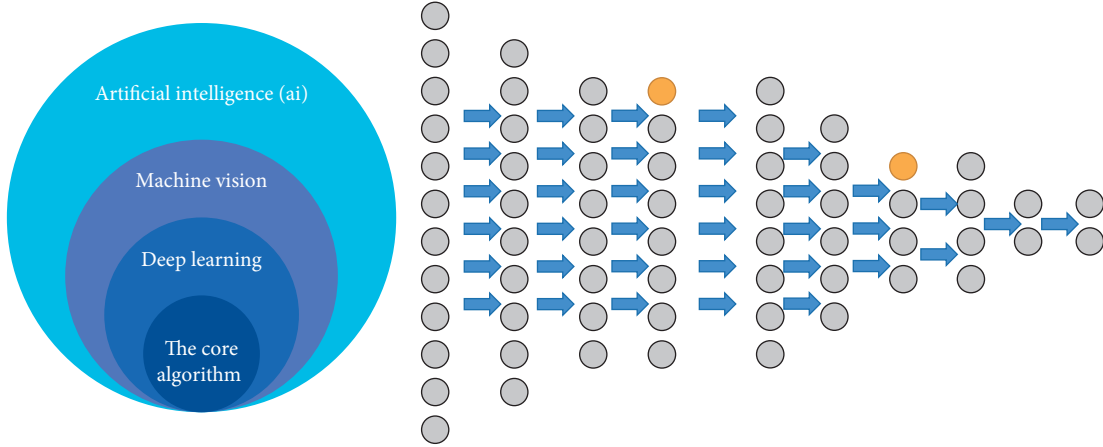
Figure 1: Multilayer neural network of deep learning technology.

achieved good results in various image recognition tasks in various fields, such as face recognition, fingerprint recognition, license plate recognition, and target tracking [4]. However, changes in illumination and viewing angles still present certain challenges to image recognition. In order to solve this problem, Li B. proposed a color image recognition method that combines image feature data and a deep trust network to construct image color data. Under the premise of the field, wavelet transform is used to describe the multiscale features of the image, and finally, the deep trust network is trained unsupervised, but the recognition rate needs to be further improved [5]. Rao M. et al. proposed an image recognition method based on the CNN-GRNN model and designed a new image recognition model that extracts multilayer features in the image through a convolutional network and uses a generalized regression neural network to replace the reverse spread the neural network to improve the generalization ability and robustness of the classifier [6]. Clabaut, T. fused Bayesian networks and artificial neural networks to create a model. The fusion model can well reflect the correlation of road traffic flow in time and space. The study verifies the prediction results of the fusion model by using the traffic flow data collected by large floating vehicles installed on the Roman road network. The verification results show that the spatial structure of the Bayesian network is effective in road traffic flow prediction under general conditions. In a few cases of nonrecurring congestion, it is more valuable to use a single-dimensional time series method to process data [7].

On the current basis, this paper proposes a kind of information fusion processing based on the combination of cloud services and big data. First, frequent item detection and association rule feature extraction are performed on the high-dimensional fusion data. Then, they use a convolutional neural network classifier to classify the extracted association rules. Combine the feature compression method to reduce the dimensionality of the classified output cloud service portfolio big data and use the clustering method to realize the classification and mining of the cloud service portfolio big data; the designed network structure adopts unsupervised bipartite K-means and convolutional neural networks. Compared with the traditional convolutional neural network, the combined method of (CNN) can reduce the training parameters and eliminate the problem of gradient dispersion.

2. Distributed Structure Model and Feature Extraction of Big Data

2.1. Feature Distribution Model of Cloud Service Portfolio Big Data. In order to realize the optimized mining of cloud service combination big data, firstly build the distributed data structure model of big data under the cloud service combination mode and use the quadruple G to represent the distributed storage center of cloud service combination big data, $G=(V, E, W, C)$ [8]. Assuming d is the phase space embedding dimension of the cloud service combination big data interaction, the multi-nonlinear component joint statistical method is used to reconstruct the high-dimensional feature space of the cloud service combination big data, and the clustering method is used for the adaptive classification of big data. Based on the above analysis, the overall structure model of the cloud service portfolio for big data mining is constructed, as shown in Figure 2.

There are many disturbance factors in the process of cloud service portfolio big data mining, which are time-varying and random [9]. The clustering method is used for big data information fusion, and the association rule term constraint equation is used to express the information flow model of cloud service portfolio big data, and it is expressed as

$$x_n = x(t_0 + n\Delta t) = h[z(t_0 + n\Delta t)] + \omega_n, \qquad (1)$$

where h is the distributed time series of cloud service combination big data, expressed as a function with a multidimensional data structure model and ω_n is the observation or measurement error of big data multisensor information fusion tracking. The distribution function description formula of the distribution structure model of cloud service portfolio big data is
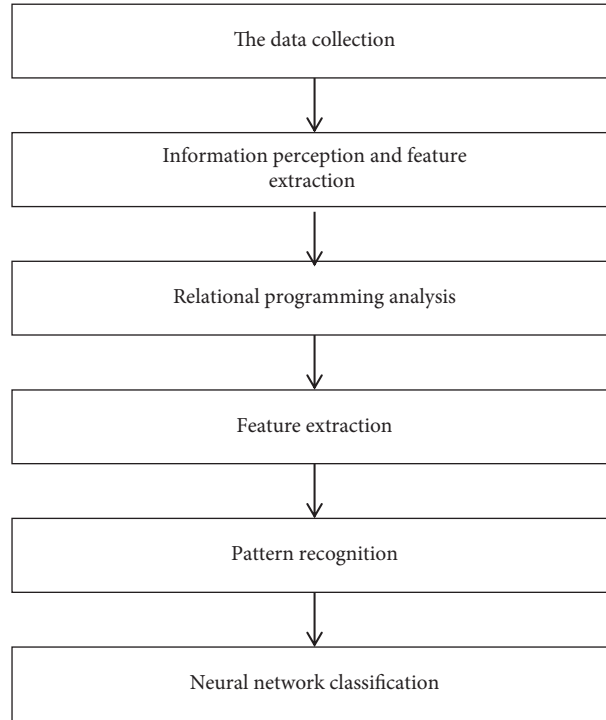
FIGURE 2: The overall structure model of cloud service portfolio big data mining.

$$X_p(u) = \begin{cases} p\sqrt{\dfrac{1-j\cot a}{2\pi}}\, e^{j(u2/2)\cot a} \displaystyle\int_{-\infty}^{+\infty} x(t)e^{j(u2/2)\cot a - jtu\csc a}\mathrm{d}t, & a \neq n\pi, \\[2mm] x(u), & a = 2n\pi, \\[2mm] x(-u), & a = (2n \pm 1)\pi, \end{cases} \tag{2}$$

where p is the order of the big data storage structure of the distributed cloud service combination and α is the time window width of statistical information sampling. Then we construct a temporal structure model of cloud service portfolio big data distribution.

The mined cloud service combination big data is reconstructed according to the five-tuple of the association rule item characteristics, the association rule knowledge base is constructed, and the characteristic identification function of the cloud service combination big data data structure is given as $P_c = \sum_{i=0}^{n}\sum_{j=0}^{n} a(i,j)P(i,j)$. The statistical regression analysis method is used to construct the nonlinear time series model of the cloud service combination big data, and the linear combination model is obtained as

$$x_k = \sum_{n=0}^{(N/2)-1} 2\left(a_n \cos\frac{2\pi kn}{N} - b_n \sin\frac{2\pi kn}{N}\right), \quad k = 0, 1, \ldots, N-1, \tag{3}$$

where a_n represents the magnitude of the cloud service portfolio big data linear programming model. For a set of continuous cloud service combination big data, the continuous template matching technology is used to analyze the distributed data structure of big data, and the information fusion processing of cloud service combination big data is combined with matching-related detection methods. The data flow processing is shown in Figure 3.

2.2. Data Feature Extraction. Suppose the amount of cloud service combination big data nodes is m, and the closed frequent itemset feature extraction output of each node is expressed. Based on the extreme learning method for the global optimization of data feature extraction, the mathematical expression of the linear programming problem for constructing cloud service portfolio big data mining is as follows:
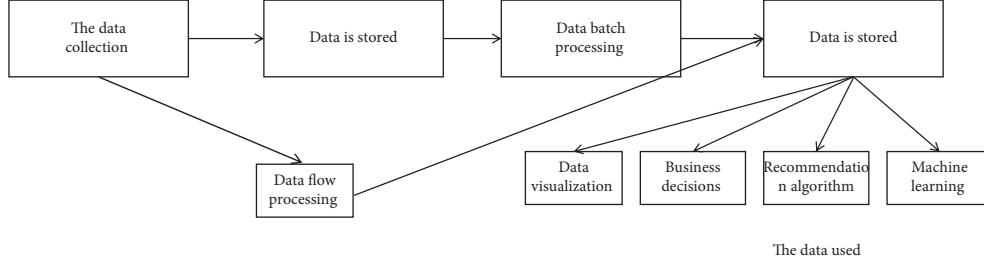
FIGURE 3: Distributed data structure of big data.

$$\min(f) = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} X_{ij},$$

$$\text{s.t} \begin{cases} \sum_{j=1}^{m} X_{ij} = a_i, & i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} X_{ij} = b_i, & j = 1, 2, \ldots, n, \\ X_{ij} \geq 0, & i = 1, 2, \ldots, m; j = 1, 2, \ldots, n. \end{cases} \tag{4}$$

Assuming that the current number of cloud service portfolio big data distribution nodes is $n, N_{1n}, \ldots, N_n$, the load of the big data to be mined in the link layer is $L_1, \ldots, L_n$, and the estimated characteristics of the cloud service portfolio big data mining output are $P_1^{\min}, \ldots, P_n^{\min}$ [10]. In the linear programming model, the feature decomposition of cloud service portfolio big data is carried out, wavelet entropy is obtained, and the information fusion processing of cloud service portfolio big data is combined with matching correlation detection methods. Perform frequent item detection on high-dimensional fusion data to realize the feature extraction of association rules, which is expressed as

$$\begin{cases} \rho_t^I = \dfrac{\sum_1^{N_I} s(i,t)}{V} = \dfrac{N_I}{V}, \\ \rho_t^R = \dfrac{\left| \sum_1^{N_R} s(i,t) \right|}{V} = \dfrac{N_R}{V}, \\ \rho_t^s = \dfrac{N_S}{V}. \end{cases} \tag{5}$$

A multivariate statistical characteristic equation is used to describe the fitting state model of cloud service portfolio big data as

Among them, N_1, N_R, and N_s, respectively, represent the average mutual information feature quantity and state distribution set of cloud service portfolio big data [11].

2.3. Data Classification and Mining Technology Optimization

2.3.1. Convolutional Neural Network Classification.
Based on the abovementioned large data distributed data structure analysis and association rule feature extraction using continuous template matching technology, the optimization design of data classification mining algorithms is carried out. This paper proposes a big data classification mining technology based on convolutional neural networks, using statistical average methods to build a regression analysis model of the big data that needs to be mined and it is expressed as

$$\begin{cases} \min \sum_{1 \leq i \leq K e \subseteq k(e)} \sum \dfrac{f(e(i))}{C(e,i)}, & 0 \leq f(e,i) \leq C(e,i), \\ & F = \text{const}, \end{cases} \qquad \sum_{1 \leq i \leq K, e \subseteq k(e)} \dfrac{f(e(i))}{C(e,i)} + \sum_{e \subseteq k(e)} \dfrac{f(e'(i))}{C(e',i)} \leq k(v). \tag{6}$$

$$\begin{pmatrix} X \\ P(X) \end{pmatrix} = \left\{ \begin{matrix} a_1, a_2, \ldots, a_m \\ p(a_1), p(a_2), \ldots, p(a_m) \end{matrix} \right\}. \tag{7}$$

Among them, $0 \le p(a_i) \le 1$ $(i = 0, 1, 2, \ldots, m)$ and $\sum_{i=1}^{m} p(a_i) = 1$ represent the autoregressive statistical characteristic parameters of cloud service portfolio big data. A convolutional neural network classifier is used to classify the attributes of the extracted association rules of cloud services combined with big data [12]. The convolutional neural network is a three-layer network structure, and the input and output iteration equation of the convolutional neural network classifier is

$$W(n+1) = W(n) - \eta \frac{\partial E}{\partial W} + \partial \Delta W(n). \tag{8}$$

Assuming that the learning step size of the convolutional neural network for big data recognition is

$$w_{sij}(n_0 + 1) = w_{sij}(n_0) - \eta_{sij} \frac{\partial J}{\partial w_{sij}}. \tag{9}$$

Using the learning algorithm of the convolutional neural network, the adaptive learning weighting coefficient of the cloud service portfolio big data classification is obtained as

$$a_{\text{desira}}^{i} = a_1 \cdot \frac{\text{Density}_i}{\sum_i \text{Density}_i} + a_2 \frac{\text{AP}_i}{\text{AP}_{\text{init}}}. \tag{10}$$

Under the constraints of $B \Rightarrow D$, $A \cap B \Rightarrow D$, and other rules, the attribute set of cloud service portfolio big data classification satisfies

$$\begin{cases} a_1 + a_2 = 1, \quad a_1, a_2 \in [0, 1], \\ \\ a_2 = \dfrac{\max_i(\text{AP}_i) - \min_i(\text{AP}_i)}{\text{AP}_{\text{init}}}. \end{cases} \tag{11}$$

The statistical quantitative set of data is $(u, v) \in E$, suppose $A \subset V$, $B \subset V$, and $A \cap B = \phi$. The convolutional neural network classifier is used for attribute classification to realize the big data reorganization and data structure rearrangement of the cloud service portfolio. The neuron structure of the convolutional neural network is shown in Figure 4.

2.3.2. Data Feature Dimensionality Reduction and Classification Mining Output.
Based on the use of convolutional neural network classifiers for attribute classification, in order to reduce the computational cost, combined with the feature compression method, the dimensionality reduction processing of the cloud service composite big data output by the classification is performed [13]. The feature compressor is described as

$$F = \overline{p}(x, y) = p(x, y) \left(\frac{v(x)}{v(y)} \right)^{1/2}. \tag{12}$$

Among them
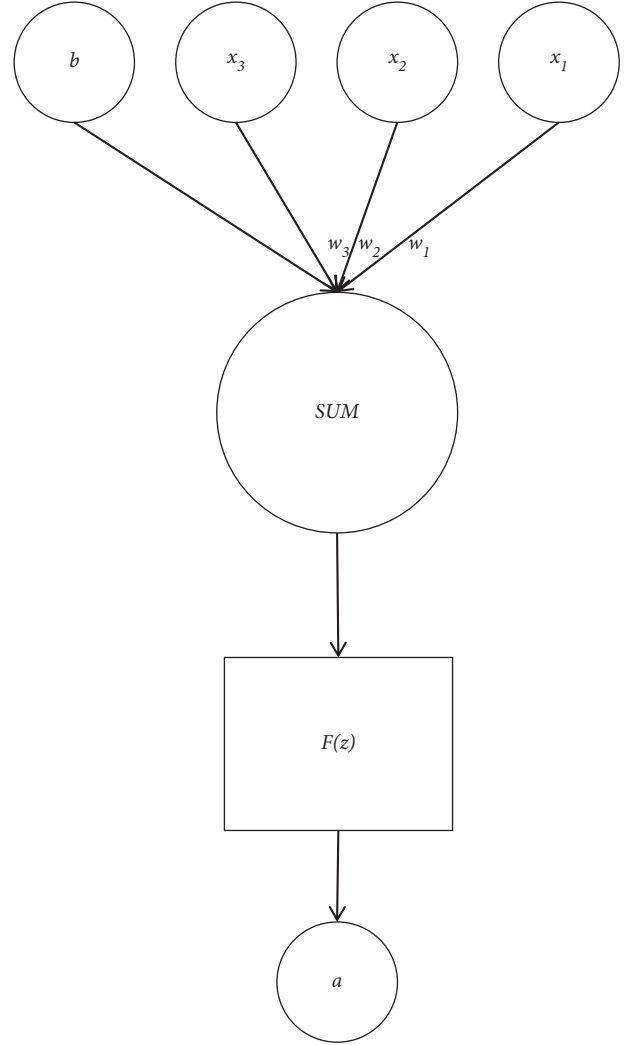
Figure 4: Convolutional neural network neuron structure.

$$p(x, y) = \frac{k(x, y)}{v(x)},$$
$$v(x) = \sum_y k(x, y). \tag{13}$$

The conceptual grid node of cloud service portfolio big data obtained by using a convolutional neural network classifier is

$$X' = \left(x'_{i0}, x'_{i1}, \ldots, x'_{i(n-1)}, y'_{i0}, y'_{i1}, \ldots, y'_{i(n-1)} \right)^T. \tag{14}$$

The dimensionality reduction of the cloud service portfolio big data output by classification is performed, and the clustering method is used to realize the classification mining of the cloud service portfolio big data [14].

3. Method for Predicting Lateral Turning Angle of Trolley Based on Bipartite K-Means Convolutional Neural Network

3.1. Trolley Lateral Control. The lateral control of the trolley is divided into image acquisition, image preprocessing, CNN model establishment, and prediction angle in this article. The overall process is shown in Figure 5.

3.2. Building a Dataset. A camera is preinstalled on the front, left, and right parts of the front end of the trolley, and then the trolley is manually controlled to avoid obstacles, and the image data of the three cameras are obtained from the upper computer. The data of the three cameras are divided into three parts: front, left, and right. At the same time, the steering gear angle is converted into a digital signal through an adjustable resistor, and then the digital signal is converted into the rotation angle of the front wheel of the trolley by the host computer [15]. In the process of data collection, in order to ensure the validity of the dataset in a variety of complex situations, the data collection should meet the following requirements:

(1) In order to avoid the influence of light, weather, and other factors on the camera, the dataset should be collected under different light and weather conditions in the same environment.

(2) In order to avoid the occurrence of too many abnormal quantities, we try to keep the angle similar to the original under the same environment, and try to avoid the appearance of the angle that deviates too much from the original [16]. When the above conditions are met, the article collected more than 400 images at each of the front, left, and right positions of the trolley. The comparison of the accuracy of data mining is shown in Figure 6.

3.3. Dataset Preprocessing. In order to ensure that the number of images is sufficient to prevent the occurrence of overfitting, this article amplifies the dataset by adding Gaussian noise and salt-and-pepper noise to the original image, and finally obtains more than 1,400 images in each of 3 locations, totaling more than 4200 image dataset, the time-domain waveform of data mining is shown in Figure 7.

Since the turning angle is converted by a digital signal, the change of the turning angle in time is not smooth. In order to smooth the data, this article selects a moving average and calculates the average with 3 time units (0.3 s) as a window. Figure 8 is a comparison diagram of the original data and the smoothed data over time.

3.4. Establishment of Convolutional Neural Network Model Based on Bipartite K-Means

3.4.1. Network Structure. This section explains the improved convolutional neural network structure for the prediction system of the lateral turning angle of the trolley. There are 1 K-means layer, 3 ordinary convolutional layers, 2 fully connected layers, and 2 dropout layers [17]. Among them, 3 convolutional layers and 2 fully connected layers are trainable layers. The height and width of the convolutional layer decrease as the depth of the network structure increases. The output node of the fully connected layer of the last layer is 1. In front of the fully connected layer, add a dropout layer to improve the generalization ability of the network and prevent overfitting [18].

3.4.2. Input Layer. The preprocessed image pixels are 160×320, and the preprocessed images at the front, left, and right positions are used as the input of the input layer. The purpose is to preserve the image information of each position of the vehicle as much as possible.

3.4.3. Dichotomous K-Means Layer. When traditional convolutional neural networks are used in image recognition, a huge dataset is required for training, and it is prone to overfitting problems. Therefore, it is thought of clustering the pictures first. The classic K-means clustering algorithm is greatly affected by the initial clustering center and it is easy to converge to the local minimum, so the bipartite K-means algorithm is used [19]. Using the bisection K-means algorithm can reduce the amount of training overhead, overcome the local optimal problem caused by the uneven distribution of the dataset, and the clustering effect is better. The images at the front, left, and right positions are passed through the bisected K-means layer in turn, and the k clustering results obtained are used as the input of the next layer. The k value of the dichotomous K-means layer designed in this paper is set as 3.

3.4.4. Convolutional Layer. The convolutional layer extracts the characteristics of the input image through convolution calculation. The network structure designed in this paper has 3 convolutional layers, and each convolutional layer decreases in turn as the network depth increases. The first convolutional layer has 16 filters, and the size of each filter is $8 \times 8 \times 3$, and the step size is 4; the second convolutional layer has 32 filters, and the size of each filter is $5 \times 5 \times 3$, and the step size is 2; the third convolutional layer has 64 filters, each filter has a size of $5 \times 5 \times 2$, and the step size is 2. There are many types of activation functions in convolution calculations. Among them, the sigmoid function is the most commonly used, and it can also be a modified linear unit (ReLu). The modified linear unit used here is to speed up the training of the network and reduce the calculation time of the network (Figure 9).

$$f(x) = \max\left(0, \sum_i w_i a_i\right). \qquad (15)$$

In the formula w is the connection weight and a is the output of the previous layer.
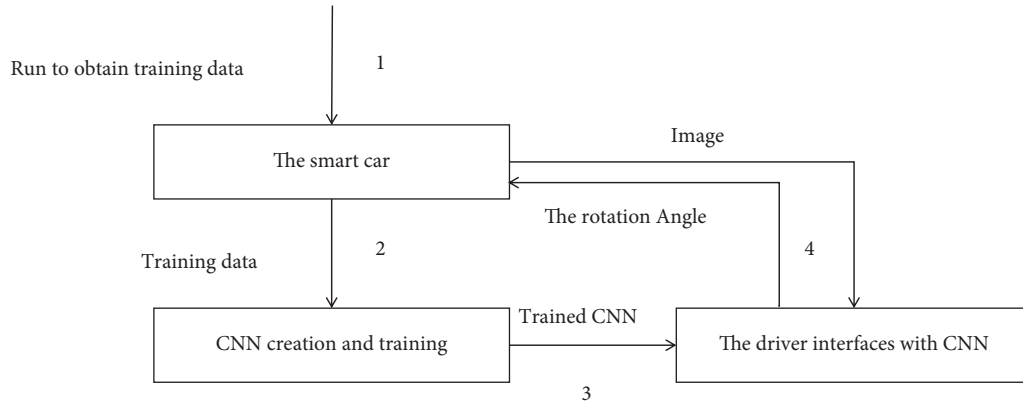
FIGURE 5: Overall flow chart.
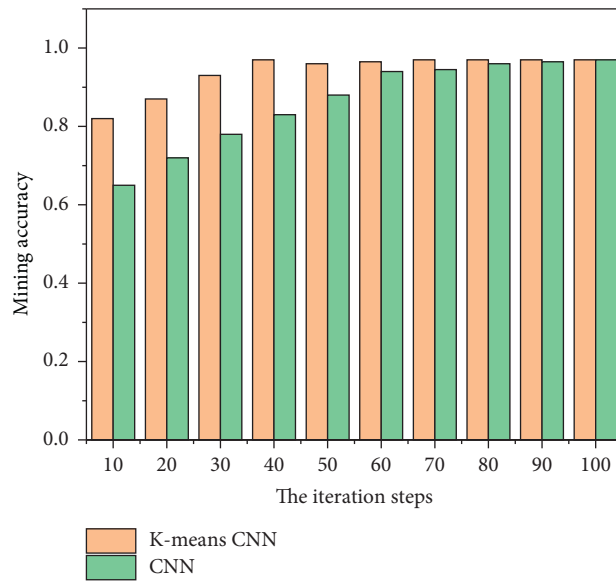
FIGURE 6: Comparison of data mining accuracy.
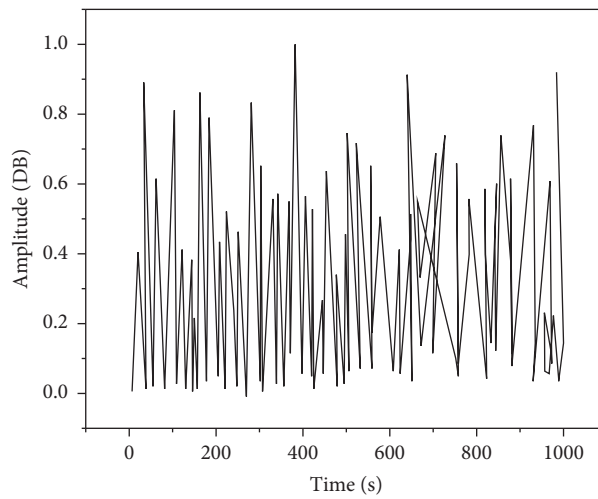
FIGURE 7: Time-domain waveform of data mining.
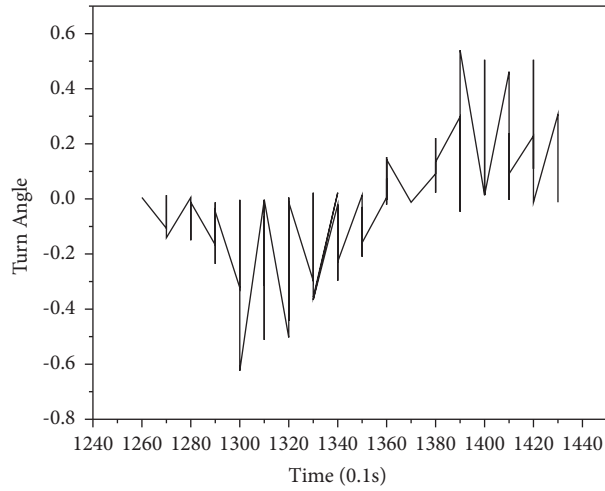
FIGURE 8: Comparison of the original rotation angle and the smoothed rotation angle.
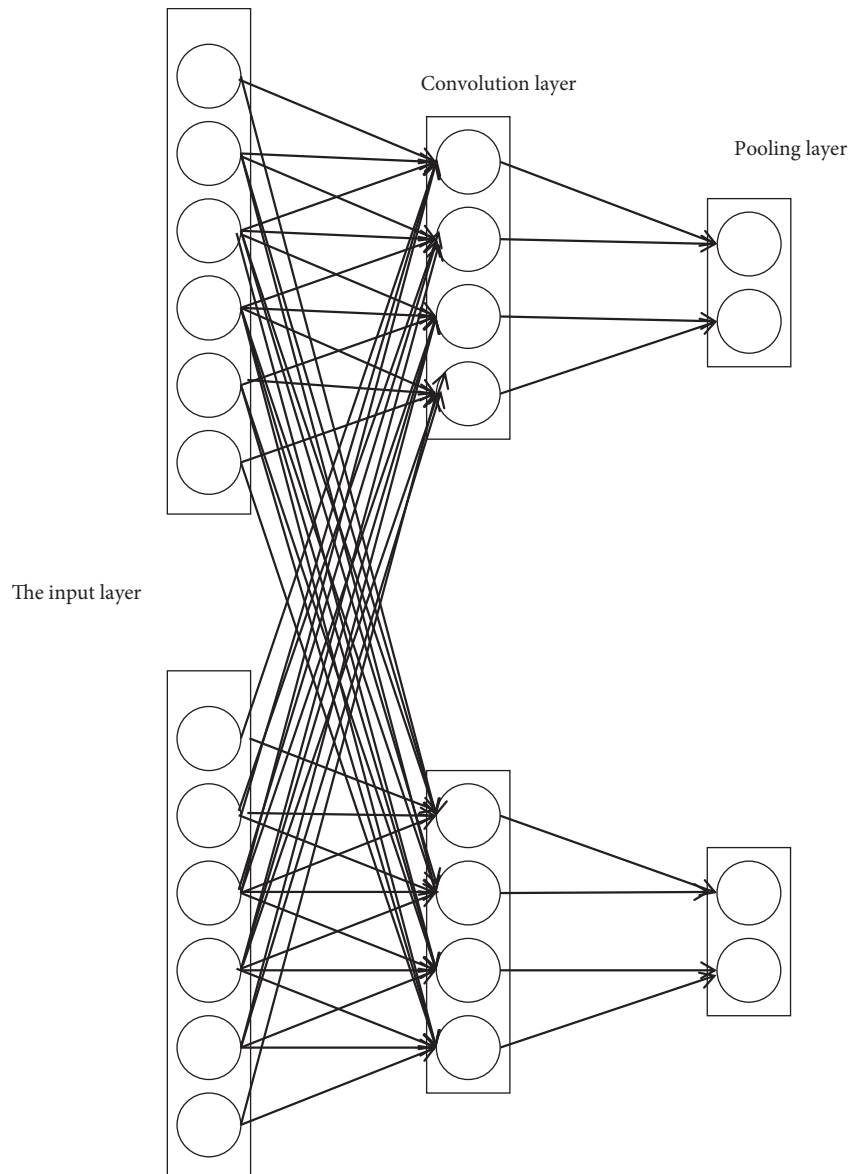
FIGURE 9: K-means convolutional neural network model.
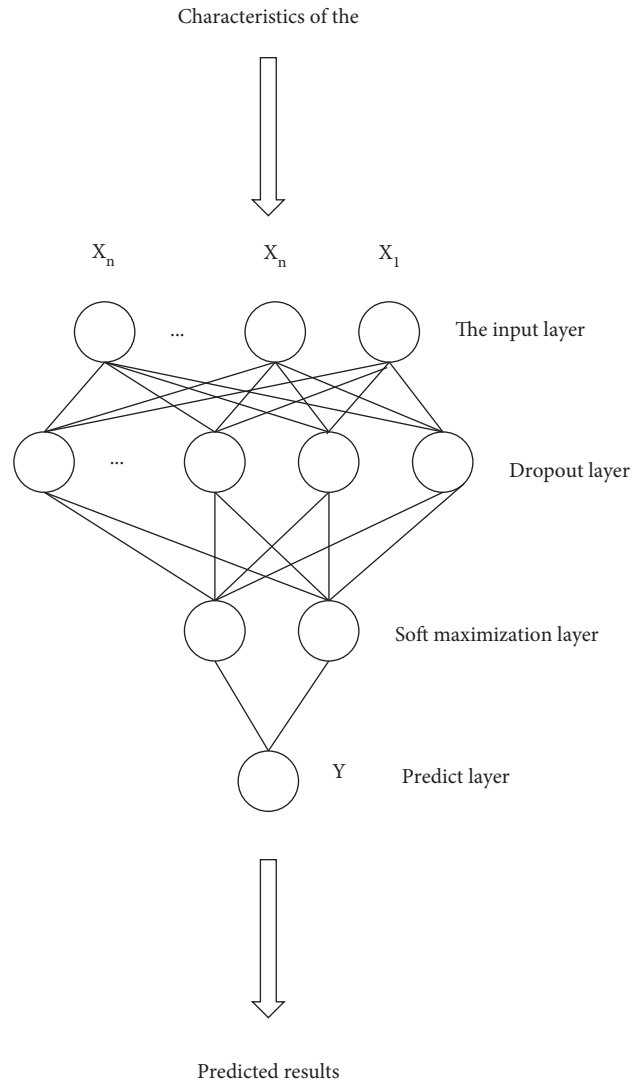
Figure 10: Classification layer network structure.
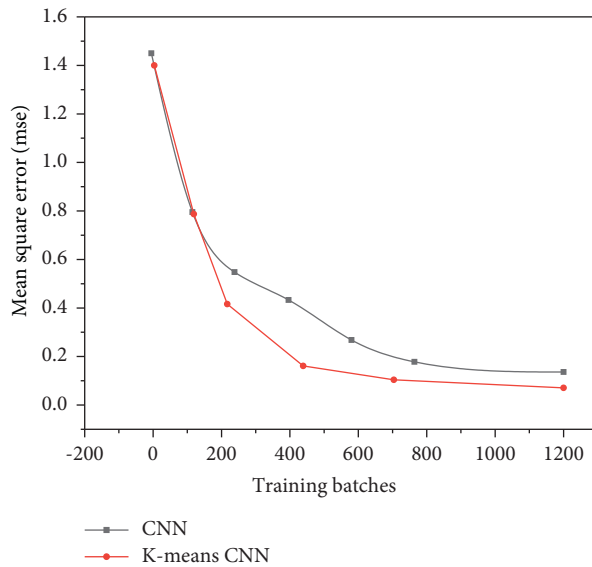
Figure 11: The first comparison of results based on the bipartite K-means convolutional neural network and the traditional convolutional neural network.
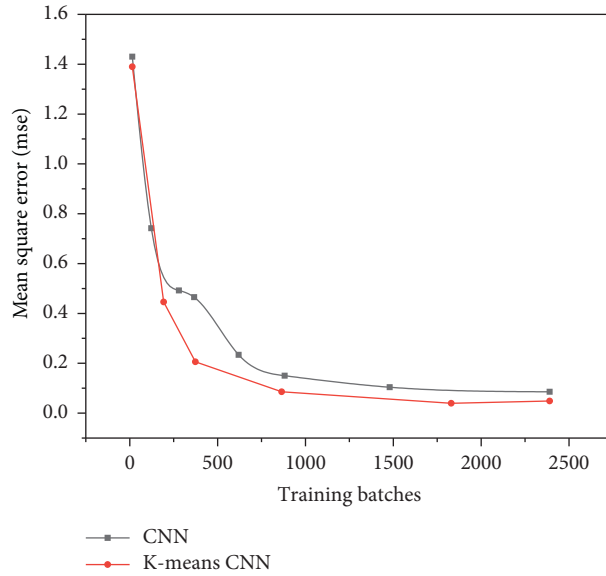
FIGURE 12: The second result comparison between the bipartite K-means convolutional neural network and the traditional convolutional neural network.
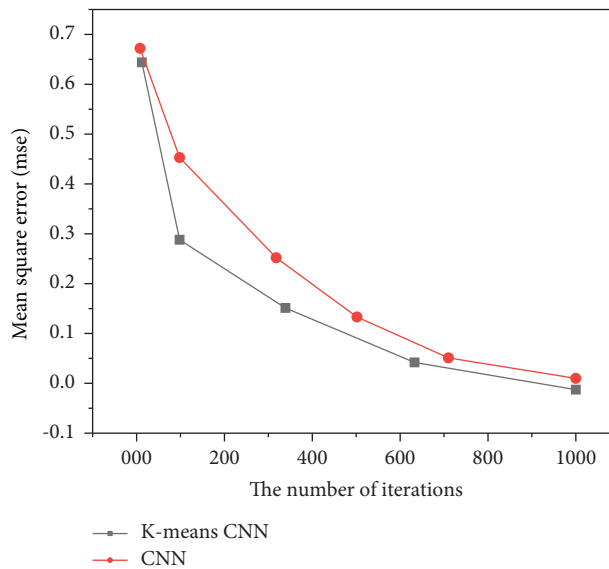
FIGURE 13: Comparison of training results based on the bipartite K-means convolutional neural network and the traditional convolutional neural network.

3.4.5. Output Layer. The output layer of the network structure is a fully connected layer, and the output node of the fully connected layer is 1. In order to avoid overfitting, a dropout layer is added in front of the fully connected layer, and the dropout rate of the dropout layer is 0.5. A dropout layer with a dropout rate of 0.2 is also added to another fully connected layer of the network structure. The network structure of the classification layer is shown in Figure 10.

4. Experimental Results and Analysis

The training results of the convolutional neural network based on dichotomous K-means and the traditional convolutional neural network are shown in Figures 11–13, and the comparison results of accuracy and error results are shown in Figures 14 and 15.

It can be seen from Figure 13 that although the two networks eventually tend to converge, if only the convolutional neural network is used for angle prediction, more iterations are needed to gradually converge. If a bipartite K-means convolutional network is added, the image category can be determined faster, so that fewer iterations are required.

It can be seen from Figures 14 and 15 that the recognition rate of the two network structures increases with the increase of the number of test pictures.
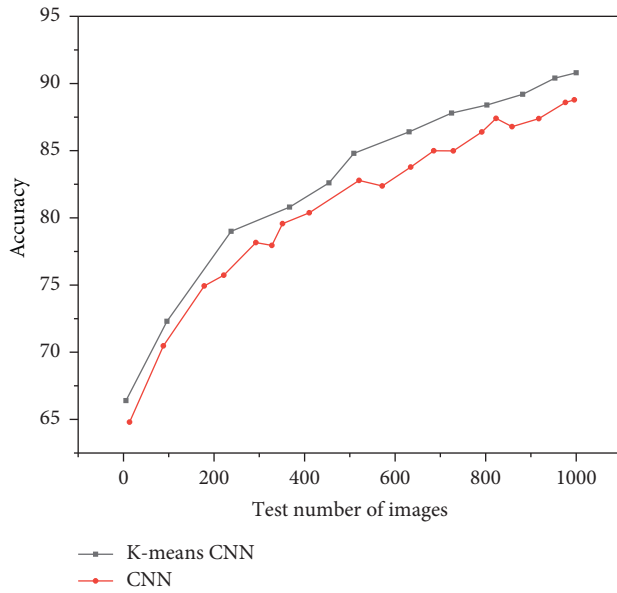
FIGURE 14: Comparison of accuracy based on the bipartite K-means convolutional neural network and the traditional convolutional neural network.
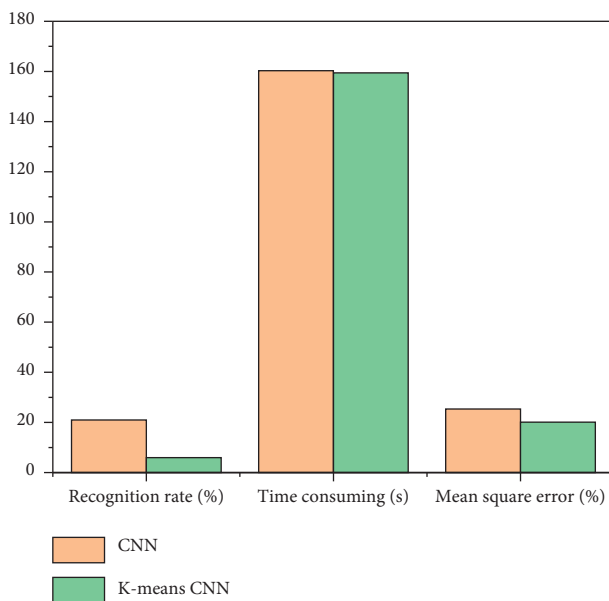
FIGURE 15: Error comparison between the bipartite K-means convolutional neural network and the traditional convolutional neural network.

5. Conclusion

A convolutional neural network classifier is used for attribute classification, combined with a feature compression method to reduce the dimensionality of the cloud service portfolio big data output by classification, and a fuzzy clustering method is used to realize the classification and mining of cloud service portfolio big data. The bisection K-means clustering method is added to the traditional convolutional neural network for optimization, and it is

verified on the image data collected by the smart car camera. The improved convolutional neural network uses the binary K-means clustering method to perform clustering learning first, so that the convolutional neural network can obtain richer input information, thereby reducing the training parameters of the convolutional neural network, accelerating the training speed of the network, and improving training accuracy. The experimental results show that the improved convolutional neural network model has faster training speed and higher accuracy than the traditional convolutional neural network.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] B. Yang, R. Liu, and E. Zio, "Remaining useful life prediction based on a double-convolutional neural network architecture," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9521–9530, 2019.

[2] Z. Liu, H. Chen, Z. Ren et al., "Deep learning audio magnetotellurics inversion using residual-based deep convolution neural network," *Journal of Applied Geophysics*, vol. 188, no. 6433, Article ID 104309, 2021.

[3] A. Viebke, S. Memeti, S. Pllana, and A. Abraham, "Chaos: a parallelization scheme for training convolutional neural networks on intel xeon phi," *The Journal of Supercomputing*, vol. 75, no. 1, pp. 197–227, 2019.

[4] H. Guan and X. Zhao, "Study on the prediction system of shrimp field distribution in the east China sea based on big data analysis of fishing trajectories," *Journal of Ocean University of China*, vol. 20, no. 1, pp. 228–234, 2021.

[5] B. Li, B. Xiao, and Y. Yang, "Strengthen credit scoring system of small and micro businesses with soft information: analysis and comparison based on neural network models," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 1, pp. 1–18, 2021.

[6] M. Rao and D. Ramesh, "Parallel cnn based big data visualization for traffic monitoring," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 1, pp. 1–13, 2020.

[7] É. Clabaut, M. Lemelin, M. Germain, M.-C. Williamson, and É. Brassard, "A deep learning approach to the detection of gossans in the canadian arctic," *Remote Sensing*, vol. 12, no. 19, p. 3123, 2020.

[8] Z. Guo, J. Zhao, Z. You, Y. Li, S. Zhang, and Y. Chen, "Prediction of coalbed methane production based on deep learning," *Energy*, vol. 230, no. 2, p. 120847, 2021.

[9] S. Tang, S. Yuan, Y. Zhu, and G. Li, "An integrated deep learning method towards fault diagnosis of hydraulic axial piston pump," *Sensors*, vol. 20, no. 22, pp. 1–20, 2020.

[10] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-lstm: convolutional lstm model for sentiment analysis in social big

data," *Information Processing & Management*, vol. 58, no. 1, Article ID 102435, 2021.

[11] J. Y. Kim and S. B. Cho, "Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network," *Expert Systems*, vol. 36, no. 4, pp. e12403.1–e12403.12, 2019.

[12] J. Dyson, A. Mancini, E. Frontoni, and P. Zingaretti, "Deep learning for soil and crop segmentation from remotely sensed data," *Remote Sensing*, vol. 11, no. 16, p. 1859, 2019.

[13] Y. Kim, P. Wang, and L. Mihaylova, "Scalable learning with a structural recurrent neural network for short-term traffic prediction," *IEEE Sensors Journal*, vol. 19, no. 23, pp. 11359–11366, 2019.

[14] C. Wang, H. Sun, R. Zhao, and X. Cao, "Research on bearing fault diagnosis method based on an adaptive anti-noise network under long time series," *Sensors*, vol. 20, no. 24, p. 7031, 2020.

[15] M.-J. Hsu, Y.-H. Chien, W.-Y. Wang, and C.-C. Hsu, "A convolutional fuzzy neural network architecture for object classification with small training database," *International Journal of Fuzzy Systems*, vol. 22, no. 1, pp. 1–10, 2020.

[16] W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu, "Deep learning methods for heart sounds classification: a systematic review," *Entropy*, vol. 23, no. 6, p. 667, 2021.

[17] S. Yin, Y. Zhang, and S. Karim, "Region search based on hybrid convolutional neural network in optical remote sensing images," *International Journal of Distributed Sensor Networks*, vol. 15, no. 5, 2019.

[18] Y. Wang, J. Yan, Z. Yang, T. Liu, Y. Zhao, and J. Li, "Partial discharge pattern recognition of gas-insulated switchgear via a light-scale convolutional neural network," *Energies*, vol. 12, no. 24, p. 4674, 2019.

[19] S. Ranjbar, K. W. Singleton, P. R. Jackson et al., "A deep convolutional neural network for annotation of magnetic resonance imaging sequence type," *Journal of Digital Imaging*, vol. 33, no. 2, pp. 439–446, 2020.

Research Article

Analysis of Bank Credit Risk Evaluation Model Based on BP Neural Network

Xiaogang Wang (ID)

College of Management, Henan University of Technology, Zhengzhou, Henan 450052, China

Correspondence should be addressed to Xiaogang Wang; xgwang@haut.edu.cn

Commercial banks are of great value to social and economic development. Therefore, how to accurately evaluate their credit risk and establish a credit risk prevention system has important theoretical and practical significance. This paper combines BP neural network with a mutation genetic algorithm, focuses on the credit risk assessment of commercial banks, applies neural network as the main modeling tool of the credit risk assessment of commercial banks, and uses the mutation genetic algorithm to optimize the main parameter combination of neural network, so as to give better play to the efficiency of neural network. After verification of various evaluation models, the accuracy of the evaluation model designed in this paper is more than 65%, while the acceptability of the evaluation results optimized by the mutation genetic algorithm is more than 85%. Compared with the accuracy of about 50% of the traditional credit scoring method, the accuracy of the credit risk evaluation using neural network technology is improved by more than 10%. It is proved that the performance of the optimized algorithm is better than that of the traditional neural network algorithm. It has important theoretical and practical significance for the establishment of the credit risk prevention system of commercial banks.

1. Introduction

The credit risk assessment has developed for a long time, and its measurement method is also changing, and the model method has been adopted for the credit risk assessment. The evaluation and management of credit risk is of more important practical significance for commercial banks [1]. The credit risk evaluation of commercial banks is a typical fuzzy problem because there are not only many factors affecting credit risk but also one factor often has multiple levels. Therefore, it is necessary to explore the index system, evaluation method, and practical application of comprehensive evaluation of credit risk. Therefore, it will be of great significance to accelerate the research work of the credit risk assessment, find the credit risk assessment method and model suitable for the actual situation of commercial banks as soon as possible, provide the basis for loan decision making of commercial banks, and finally improve the competitiveness of commercial banks [2]. To solve the

problem of the credit risk assessment, neural network technology has its unique advantages in which the most important is the nonlinear mapping ability of neural network. Through the forward neural network, the nonlinear relationship between credit index and rating can be described quantitatively, so as to meet the classification of customer credit rating according to index data [3].

Based on the in-depth study of the current situation and application characteristics of the credit risk assessment of commercial banks, combined with the characteristics of data mining and neural network technology, this paper puts forward a credit risk assessment system [4], improves and verifies the proposed model, and provides a feasible solution for the credit risk assessment. In the assessment system, the core task of data mining is credit risk classification. There are also many data mining methods that can realize classification, such as decision tree and neural network. According to the application characteristics and data characteristics of the credit risk assessment, neural network has better pertinence

and applicability than other classification methods. Therefore, neural network method is used to undertake the core task of data mining [5], namely, credit risk classification. However, the collected data cannot be directly used to train the neural network classification model, but must go through certain preprocessing and pattern division. Preprocessing includes supervised discretization, which is to determine split points in a way that maximizes the purity of the interval. However, in practice, this approach may require arbitrary determination of the bin purity and minimized bin size. To address this, some statistics-based methods divide the interval with each attribute value and create larger intervals by merging adjacent intervals similar to those based on statistical tests. The outlier analysis function of data mining is used to find the samples containing outlier data in the training samples, and discard or flatten these samples so that the neural network model can more truly reflect the mapping relationship between credit indicators and credit ratings [6]. In fact, pattern division is also classification. It distinguishes enterprise customers according to their financial data so that the financial data of the same kind of enterprise customers fall within a reasonable range, to improve the classification accuracy of the neural network model [7]. Although BP algorithm has been widely used, the main disadvantage is that the training process is uncertain. The genetic algorithm is a global optimization algorithm. This algorithm is essentially different from the single-point search method. It adopts the method of processing multiple individuals in the search space at the same time. It has good global search performance and effectively reduces the possibility of falling into the local optimal solution [8]. Therefore, the emergence of the genetic algorithm makes the training of neural network have a new look. Using the genetic algorithm instead of BP algorithm to search the connection weight of neural network is expected to solve the problem of BP network falling into local minimum. To sum up, the credit risk assessment system studied in this paper uses neural network to complete the function of the credit risk assessment [9].

In this paper, genetic algorithms are introduced to optimize the network established by backpropagation algorithm to form a hybrid algorithm. For the credit risk, a backpropagation network includes input layer, output layer, and several hidden layers. The principle of the network for the credit risk assessment of commercial banks is to take the index information used to measure the financial and non-financial status of loan enterprises as the input vector of God and network, the classification result is used as the output vector of neural network, and the network is trained with training samples to obtain different output values. After the training, it can be used as an effective tool for the credit risk assessment and prediction. After the training, credit risk can be effectively predicted. This paper compares and analyzes the prediction results of neural network optimized by the mutation genetic algorithm to verify its application in the credit risk evaluation of commercial banks. The experimental results show that the mutation genetic algorithm has the ability to optimize the training of neural network, and it is feasible and effective to apply the neural network model optimized by the mutation genetic algorithm to the credit risk evaluation of commercial banks.

2. Related Work

The credit risk assessment is the core content of bank risk management. Credit risk will lead to bank bankruptcy. Therefore, since the birth of the bank, banking experts have begun to explore the credit risk assessment [10]. The evaluation of creditability involves many factors, generally including the following five independent and related factors, namely, economic strength, growth ability, asset management ability, solvency ability, and profitability. Since the 1980s, the evaluation of credit risk has changed from relying on the experience and subjective analysis of banking experts to a credit scoring model based on financial indicators, mainly including linear probability model, logit model, probit model, and discriminant analysis model [11]. In the late 1980s, some experts at home and abroad began to use artificial intelligence theory and neural network theory to improve the accuracy and accuracy of distinction. The application of these methods overcame the requirements of the above traditional statistical methods for assumptions and the shortcomings of statically reflecting credit risk. Since the 1990s, the loan profit of commercial banks has continued to decline and the risk of off-balance sheet business has increased, which has prompted banks in various countries to pay more attention to the management and prevention of credit risk. Many Western banks have begun to explore the use of modern financial theory and mathematical tools to develop new credit risk measurement models [12]. They began to develop new and more complex credit evaluation systems, changing from simple single loan risk to measuring centralized risk. All these efforts have made great progress in the measurement and control technology of credit risk in the Western banking industry and produced a series of successful quantitative management models of credit risk. Comprehensive models can be divided into two categories: one is the combination theory model, such as JPMorgan's Credit Metrics and KMV method [13] and the other is the default model, such as credit risk + of Credit Suisse and portfolio view model of McKinsey. The biggest difference between the two models is that the default model estimates the future default distribution of the portfolio based on theoretical analysis, while the portfolio theoretical model estimates the future default distribution of the portfolio from historical data [14].

In the research of the credit risk identification, the most commonly used neural networks are BP neural network [15], RBF neural network, and probabilistic neural network. BP network is a one-way propagation multilayer feedforward network [16]. Due to its good mapping ability to approximate nonlinear functions and parallel processing, it can approximate any nonlinear mapping relationship to obtain the solution of the problem; instead of relying on the prior knowledge and rules of the problem, it has strong adaptability and is very suitable for solving the classification and pattern recognition problems of irregular, multiconstraint, or incomplete data [17, 18]. Due to the perfect theoretical

basis and few restrictions that can solve the characteristics of data samples, neural network is widely favored in solving classification problems, especially the test data generation based on genetic algorithm. This method first converts the structural coverage problem into a numerical function optimization problem and then generates the expected test data [19]. The genetic algorithm (GA) was first proposed by American scholar Holland in 1975. It is a computational model simulating Darwin's genetic selection and natural elimination of biological evolution. Its outstanding feature is that it contains steps very similar to biological genetics and evolution [20]. The main advantages of GA are simple, universal, strong robustness, suitable for parallel distributed processing and wide application range [21]. The search performance of genetic algorithm is mainly realized by three operators: selection, crossover, and mutation, especially the mutation operator. It can expand the search range of the algorithm and is more likely to find the global optimal solution. However, due to the wide search range, the convergence speed is affected, so it is very important to design adaptive mutation operation. The traditional genetic algorithm and various improved algorithms use the method of randomly selecting a bit to flip the variable value in the mutation operation, which cannot effectively improve the search performance and stability of binary coding. Variation testing was first proposed by Demilo and Hamlet et al. It is a software testing technology to measure the effectiveness and completeness of test case set [22]. In order to carry out variation test, first, make grammatical small changes to a certain statement of the program to produce a new program, which is called a variant.

3. Neural Network Algorithm Optimized by Mutation Genetic Algorithm

3.1. Network Structure of Evaluation Model. Neurons and the relationship between them define the structure of neural networks. For the credit risk assessment model, its structure can be expressed by the number of input and output nodes, the number of hidden layers, and the number of nodes in each hidden layer of backpropagation network. In the network structure of the evaluation model, the number of input nodes can be obtained intuitively, which is the number of credit risk evaluation indicators. The number of output nodes of the model can be one or multiple. For the classification model, the number of output nodes is related to the number of classification categories. Assuming that the credit rating is m, the number of output nodes of the evaluation model can be m or $\log_2 m$. In most cases, in order to simplify the structure and improve the training efficiency, a network model with multiple outputs is generally transformed into multiple network models with one output [23]. The structure of each model is the same, but each part is given to a separate model to ensure that the same weights are not applied to the entire image. This paper avoids full weight assignment and keeps weight sharing local; then, each model performs convolution and max pooling and generates some kind of output that outputs a dense layer. Based on this principle, this paper simplifies the output of the evaluation model. Firstly, according to the business needs, there are five levels

of credit rating: $A++$, $A+$, A, $B+$, and B; secondly, referring to the credit scoring method, the output of the evaluation model is converted into a continuous variable. Different value ranges of variables correspond to different credit grades, as shown in Table 1.

In order to evaluate the model, it is very important to determine the reasonable number of hidden layers and the number of nodes in each hidden layer. It is generally believed that increasing the number of hidden layers can reduce the network error and improve the classification accuracy of the evaluation model, but it also makes the model complex, the training time is too long, and it tends to "overfit."

The error bound of the credit risk assessment model is calculated according to Table 1. Taking class $A++$ credit rating as an example, when calculating the model output error, $A++$ takes 0.95, while when converting the assessment model output to credit rating classification, the value of $(0.9, 1.0)$ interval is mapped to class $A++$, and the maximum difference is used to calculate the output error of the model, and the following can be obtained

$$E = \frac{1}{2} \sum_{k=1}^{l} (d_k - y_k)^2. \tag{1}$$

Considering the convergence speed and classification accuracy of the evaluation model, $E^P = (1/2)(0.85 - 0.8)^2 = 0.00125$, the error bound E of the evaluation model is 0.001. In the process of network training, the error threshold should be predetermined according to the actual situation. The selection of the error threshold is completely determined by the convergence speed of the network model and the learning accuracy of the specific sample. When the E value is selected to be small, the learning effect is good, but the convergence speed is slow and the training times increases. The opposite is true if the E value is larger.

Generally, when the neural network model enters the later stage of training, the more stable is the connection weight between neurons; at this time, the learning rate should tend to be smaller, because the larger learning rate is easy to cause oscillation in the modification process of weight W. At the beginning of training, in order to speed up the convergence of the network, the learning rate is often adjusted to a large value, which needs to change the learning rate dynamically in the learning process. The standard for adjusting the learning rate is whether the error function can be reduced, and the adjustment formula is as follows:

$$\eta_{t+1} = \begin{cases} \alpha\eta_t, & \alpha > 1, \\ \beta\eta_t, & \beta < 1. \end{cases} \tag{2}$$

Among them, η is the learning rate; t is the training times; α, β, and k are ratio factors; and E is the error function,

$$\eta_t = \begin{cases} \alpha\eta_{t-1}, & E_t < E_{t-1}, \\ \beta\eta_{t-1}, & E_t > kE_{t-1}, \\ \eta_t, & \text{others}, \end{cases} \tag{3}$$

and the values in this paper are as follows: $\alpha = 1.05$, $\beta = 0.7$, and $k = 1.04$.

Table 1: Relationship of output of evaluation model and credit rating.

Credit rating	Output of training phase evaluation model
$A{+}{+}$	$(0.9, 1.0]$
$A{+}$	$(0.8, 0.9]$
A	$(0.7, 0.8]$
$B{+}$	$(0.6, 0.7]$
B	$(0, 0.6]$

In order to avoid model training trapped in shallow local minima, momentum term can be considered [24]. The weighted adjustment formula with momentum term is

$$w(t+1) = w(t) + \frac{\eta \cdot \partial E}{\partial W_t} + \alpha^2 \Delta(t), \qquad (4)$$

where α is the momentum factor, and its value is discussed in the previous section of this paper. The effect of introducing momentum term is to make the η value change equivalently instead of being a constant value in the learning process, to make the adjustment change toward the average direction of the bottom without large oscillation; that is, momentum plays the role of buffer and smoothing, and finally speeds up the learning speed of the evaluation model.

The momentum adding method is suitable for batch learning. In this way, the error is the sum of the output errors of all training samples, that is,

$$E = \frac{1}{2} \sum_{p=1}^{P} \sum_{k=1}^{l} \left(d_{p,k}^2 - y_{p,k}^2 \right). \qquad (5)$$

Moreover, the error backpropagation is started after the learning of all training samples; that is, the connection weights of each layer are updated. Error backpropagation is that the output error is transmitted back to the input layer by layer through the hidden layer in some form, and the error is allocated to all units of each layer, so as to obtain the error signal of each layer, which is used as the basis for correcting the weight of the unit. The credit risk assessment model optimized adopts the mutation genetic algorithm.

3.2. Neural Network Optimized by Mutation Genetic Algorithm.

The credit risk assessment system is the core part of credit management information system and the basis for other credit businesses. Its main function is to realize the bank's internal credit rating and provide the basis for other credit businesses and credit risk management. The general structure of the credit risk assessment system consists of three components:

(1) Input. Input the index data of credit rating. After preprocessing, the index data are used by the evaluation system. The selection of credit rating indicators plays a very important role in the credit risk assessment, which will be discussed below.

(2) Evaluation model. The credit risk assessment model is the core part of the credit risk assessment system. The assessment model reflects the mapping relationship between credit assessment indicators and credit rating, and this relationship can be self-adjusted with the change of environment.

(3) Output. From a business perspective, the output is the result of credit rating, that is, credit rating, such as $A{+}{+}$, $A{+}$, or $B{+}$, which can be summarized as classification results. From the perspective of the evaluation model, the output may be a continuous value, which can be converted to the credit rating required by the business after a certain mapping relationship.

Normalization of indicator data refers to scaling the indicator data to make it fall into a small specific interval [25]. The general iterative method is easy to fall into the trap of local minima and the phenomenon of "infinite loop" occurs, which makes the iteration impossible. The genetic algorithm overcomes this shortcoming well and is a global optimization algorithm. Normalization is particularly useful for classification algorithms, such as algorithms involving neural networks or distance metric classification algorithms such as nearest neighbor classification and clustering. For the credit risk assessment model using BP algorithm, standardizing the input values of training samples will help to speed up the learning speed of the assessment model.

The order of the mutation genetic algorithm is selection, crossover, and mutation. We adopt adaptive crossover and mutation probability, as shown in formulas (5) and (6), respectively:

$$p_c(g) = \frac{p_c(g-1) - [p_c(0) - 0.25]}{\max}, \qquad (6)$$

$$p_m(g) = \frac{p_m(g-1) - [0.25 - p_m(0)]}{\max}, \qquad (7)$$

where g represents the generation algebra, and max is the maximum number of generations.

The basic flow of the mutation genetic algorithm is shown in Figure 1. Similar to the biological evolution process in nature, the operation process of the genetic algorithm is also a repeated iterative process. The population of generation t is recorded as $P(t)$. After one generation of mutation, inheritance, and evolution, the population of generation $t+1$ is obtained. They are also a population composed of m individuals, recorded as $P(t+1)$. This population continues to undergo mutation, inheritance, and evolutionary operation, and each time, according to the rules of survival of the fittest, more individuals with higher fitness values are inherited to the next generation. In this way, a good individual c^* will be obtained in the group, and its corresponding phenotype x will reach or close to the optimal solution x^* of the problem.

3.3. Implementation of BP Neural Network Algorithm Optimized by Mutation Genetic Algorithm.

According to the theoretical analysis, we implement the process of mutation genetic neural network algorithm according to the following steps:
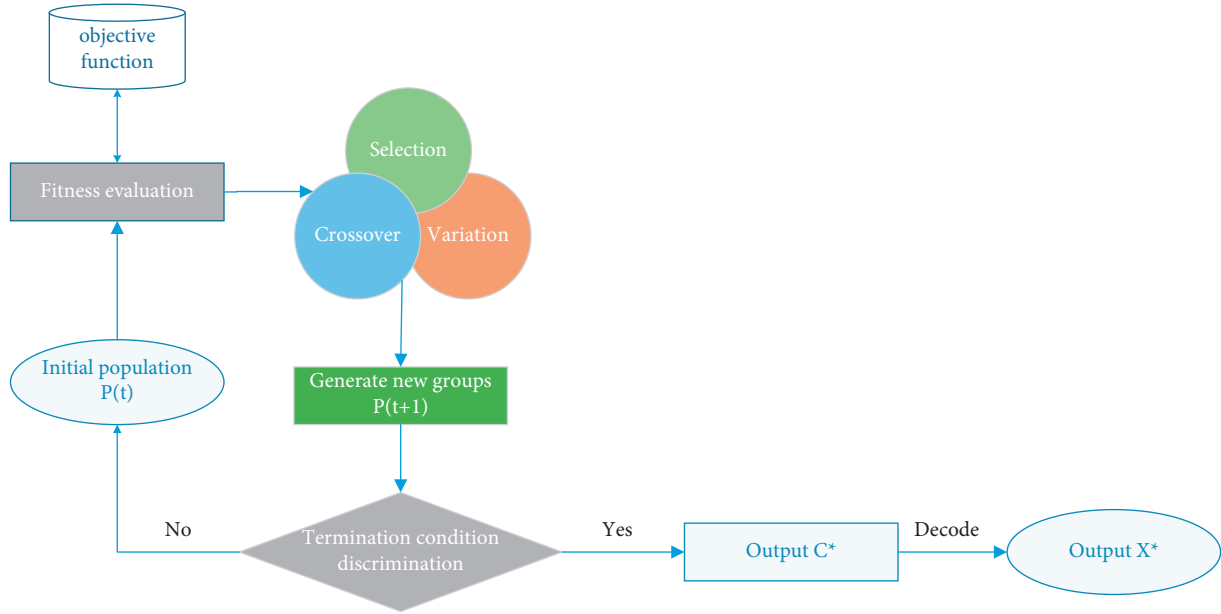
FIGURE 1: The basic flow of the genetic algorithm.

(1) Build hidden layer

 (i) It does not have a fixed formula to apply. It is usually to get a more appropriate value from experience and multiple tests. Through the repeated tests of the test samples in the previous section, the number of nodes in the hidden layer is set at 16.

(2) Coding

 (i) In this experiment, 18 neuron nodes are set in the input layer of the variant genetic neural network as risk prediction, 6 nodes are set in the output layer, the number of hidden layer nodes is 16, and the coding length formula is as follows:

$$m_n = i \cdot m_i + j \cdot m_i, \tag{8}$$

 (ii) i is input layer, j is output layer, and mi is hidden layer.

(3) Initialization function

 (i) Here, we initialize through the initializega function. The initialization function used in this paper is initp$p$ = initializega (pop, A+, "gabpeval")

(4) Fitness function

 (i) Each chromosome in the genetic algorithm corresponds to a solution of the genetic algorithm. Generally, we use fitness function to measure the advantages and disadvantages of this solution. The formula of fitness function in this paper is

$$f = k \sum_{j=1}^{n} abs\left(y_j, x_j\right)^{-1}, \tag{9}$$

 (ii) where k is the coefficient, n is the number of network output nodes, yi is the expected output, and x_i is the predicted output of the i-th node.

(5) Set the parameters of genetic algebra of variation

 (i) Here, the genetic algebra is set at about 100. In the selection operation, choose those with a larger proportion of fitness values to the sum.

(6) Cycle step 2 to step 5 to obtain the simulation results optimized by the mutation genetic algorithm until the training goal is reached or the number of iterations reaches the upper limit goal.

3.4. Credit Risk Evaluation System Architecture of Commercial Banks. The credit risk assessment system is a data analysis system based on pattern recognition. Firstly, a specific learning algorithm is used to find out the patterned relationship between customer data and credit rating from a large amount of historical data, and then expresses and stores this relationship in the form of the evaluation model for future credit risk evaluation. Therefore, from the perspective of its function, the credit risk assessment system should include the following contents: data collection, data storage, data preprocessing and data maintenance, assessment model training, assessment model management, and assessment model application. Figure 2 shows the overall architecture of the credit risk assessment system.

Among them, evaluation model training and evaluation model application constitute the core module of the credit risk evaluation system. The main function of the evaluation model training module is to generate and train the model, and store the evaluation model in the model base.

Figure 3 shows the structure of the evaluation model training module. As can be seen from the figure, the model
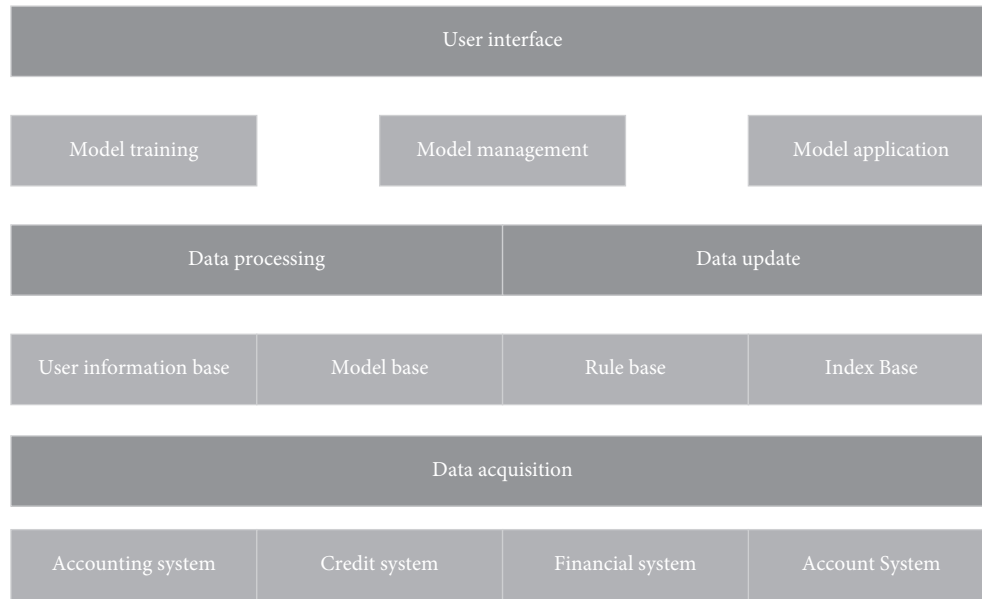
Figure 2: General framework of the credit risk assessment system.
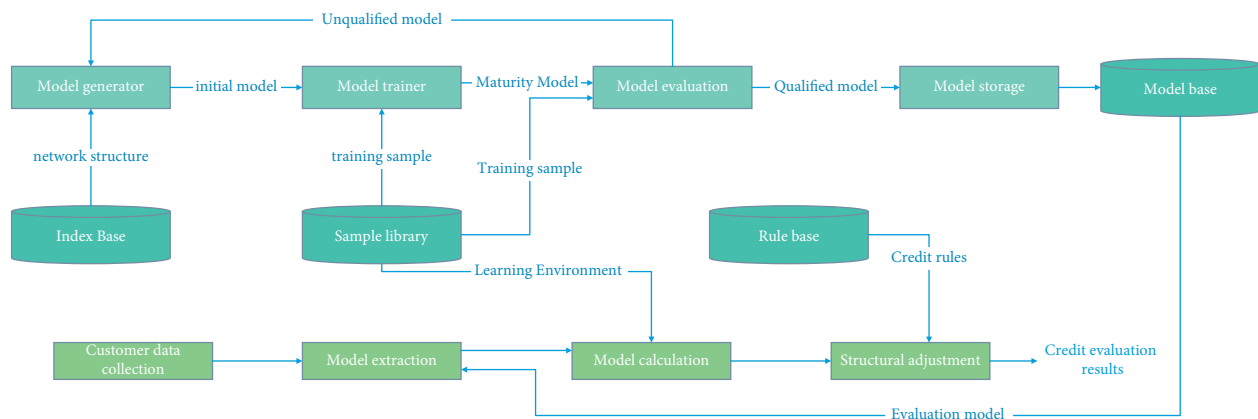
Figure 3: Structure diagram of application module of the credit risk assessment model.

training module is mainly composed of model generation, model training, and model evaluation:

(1) Model generation. Model generation first obtains the mode category of the model to be trained (generally classified by industry) from the user interface and then obtains the number of all indexes participating in the evaluation under the mode from the index library. From this number, the number of input neurons and output neurons of the neural network model can be obtained, and the number of neurons in the hidden layer can be calculated. Another main function of the model generator is to initialize the weights or thresholds of the model and learning parameters.

(2) Model training. Model training reads out the training samples of the corresponding pattern category from the sample database, processes the data, learns the processed data, and adjusts the weight and threshold of the initial model.

(3) Model evaluation. Obtain test samples from the sample library, test the trained evaluation model, and check the evaluation accuracy of the evaluation model. If the accuracy is greater than the accuracy set by the system, the model will be put into storage. Otherwise, discard the model and start training again. The first mock examination is repeated. Sometimes, it is necessary to train the evaluation model of the same pattern type for many times to get the highest accuracy evaluation model.

4. Experiment and Analysis

4.1. Preparation of Sample Data. A total of 360 data samples were collected in this paper, which were distributed in five industries. After the mode division in the previous section, 360 samples are divided into 7 credit risk assessment modes, and each mode contains more than 30 samples. This section takes industrial enterprises as an example to introduce sample data preparation and preprocessing.

Industrial enterprises include two evaluation models: small industrial enterprises and large- and medium-sized industrial enterprises. Among them, small industrial enterprises include 34 samples and large- and medium-sized industrial enterprises include 45 samples. Figure 4 shows the distribution of credit rating classification in the sample.

The sample data consist of training samples and test samples. The test samples are randomly selected from the overall samples, generally accounting for about 20% of the total samples. In order to ensure the rationality of the test, this paper requires that each credit rating level contains at least one test sample, and the distribution of credit rating in the test samples is similar to that of the overall samples. The distribution of test samples under the two evaluation modes is shown in Figure 5, while 0101 evaluation mode contains 7 test samples and the remaining 27 samples are used as training samples, and 0102 mode contains 9 test samples and the remaining 36 samples are used as training samples.

Before model training, sample data need to be pre-processed to improve the efficiency of model training and evaluate the accuracy of the model. The main idea of the BP algorithm is to input the learning samples and use the backpropagation algorithm to repeatedly adjust the weights and deviations of the network to make the output vector as close to the expected vector as possible. When the training is completed, the weights and biases of the network are saved. In this paper, the collected qualitative and quantitative indicators of enterprises are standardized, and the index data are transformed into dimensionless data in (0, 1) interval, so as to meet the requirements of input data of variant genetic neural network model. Data processing is divided into two parts: qualitative index processing and quantitative index processing.

4.2. Credit Risk Assessment Classification Model.

Firstly, the basis and activate function of neuron operation at each node of the evaluation model are selected. According to the general definition of variant genetic neural network model, the basis function of the evaluation model adopts linear function,

$$u = x^T w - \theta. \tag{11}$$

The activation function adopts sigmoidal function,

$$f(u) = [1 + \exp(-\lambda u)]^{-1}. \tag{12}$$

Secondly, the neural network structure of the credit risk assessment classification model is determined. According to the discussion in Section 3.4 of this paper, the evaluation model adopts the variant genetic neural network model with three-layer structure, and its structure can be expressed as $n \times m \times l$ where n is the number of input nodes, and the value is the number of evaluation indicators; l is the number of input nodes, and the value is 1; and m is the number of hidden layer nodes, and the value is related to many factors. Next, taking the evaluation model of industrial enterprises as an example, the hidden layer node is determined.
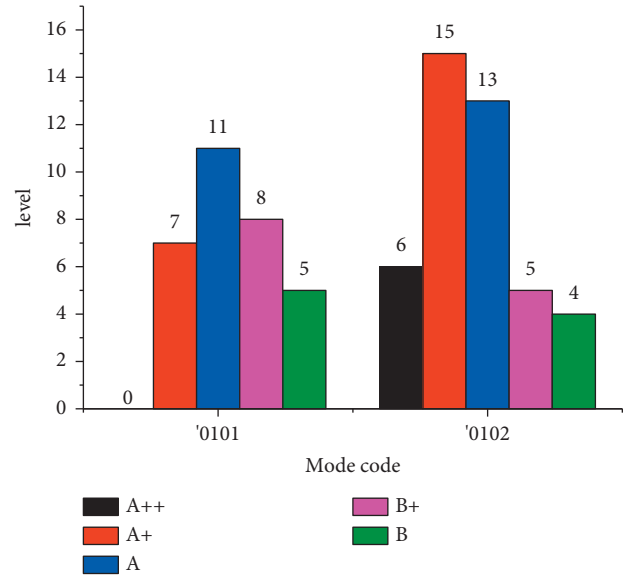
Figure 4: Classification and distribution of credit rating.
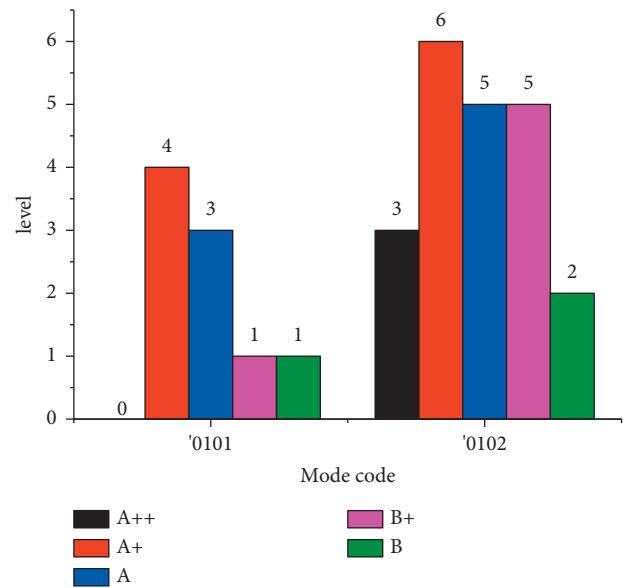
Figure 5: Classification and distribution of credit rating in test samples.

The evaluation model structure of industrial enterprise evaluation model is $28 \times m \times 1$, and the number of training samples for small- and medium-sized enterprises is 27, and the number of training samples for large- and medium-sized enterprises is 36. According to the rule that the number of hidden layer nodes is between 50% and 70% of the sum of the number of input nodes and output nodes, the value range of the number of hidden layer nodes m is about $14 < m < 21$. In order to determine the reasonable number of hidden layer nodes, this paper constructs an evaluation model for each value within the value range and determines the most reasonable number of hidden layer nodes by comprehensively comparing the training efficiency and classification accuracy of each model. Table 2 shows the training efficiency and classification accuracy corresponding

TABLE 2: Model training under a different number of hidden layer nodes.

Mode code	Number of hidden layer nodes	Simulation convergence
	5	Within 2000 training times, the simulation does not converge
	10	>1500 training,
0101	15	>1300 training,
	20	<1200 training
	25	<1100 training
	5	Within 3500 training times, the simulation does not converge
	10	Within 3500 training times, the simulation does not converge
0102	15	>1800 training
	20	<1800 training
	25	<1600 training

to the number of nodes in different hidden layers under the same other conditions.

After comparison, we finally selected a three-layer mutation genetic network structure with 16 hidden layer nodes as the evaluation model, and the specific structure is $28 \times 16 \times 1$; that is, the evaluation model includes 28 input nodes, 16 hidden layer nodes, and 1 output node. The learning and calculation process of the evaluation model, for example, K, under this mode can be obtained.

According to the analysis on the setting of learning times, target error, and learning rate in the previous chapters, after establishing the variant genetic neural network with MATLAB software, the initial parameters of these three parameters need to be set before network training on normalized samples, where net.trainparam.epochs is the number of network training, net.trainparam.goal is the network target error, and net.trainparam.lr is the network learning rate. When the initial parameters of the three are set, the train function is called to train the input training samples. The general form of the function is [net, tr] = train (net, P, T), net on the left side of brackets represents the genetic neural network before training, tr storage training records, including the number of network training and the error of each training, can generate the network error variation curve in real-time; net on the right represents the mutation genetic neural network obtained after sample training, T is the vector set of expected output of training samples, and P is the vector set of training samples.

The procedure is as follows:

(i) %Set training parameters

(ii) net.trainParam.show = 50

(iii) net.trainParam.epochs = 3000

(iv) net.trainParam.goal = le- 3; net.trainParam.Ir = 0.3

(v) %Network training

(vi) [net, tr] = train (net, p, t)

4.3. Training Process and Results. In the variant genetic neural network model, there are several very important learning parameters: the main one is the learning rate η and momentum factor α. The setting of these parameters has a great impact on the training efficiency of the credit risk assessment model. In Section 3.1, it is determined that the value of convergence error bound e is 0.001. In order to

enhance the convergence efficiency, a large learning rate η is usually used, but η too large may lead to oscillation near the stable point or even nonconvergence.

For the above situation, this paper analyzes different learning rates η and momentum factor α simulation investigation is carried out to determine the appropriate learning parameter value. Firstly, 28 are generated by MATLAB neural network toolbox $\times 18 \times 1$, and initialize, and then train the network. Set the convergence error bound $e = 0.001$, the maximum training times $T_{max} = 2000$, and select traingdm as the learning function η and α take values for network training. The network training error curve is shown in Figures 6(a) and 6(b). Figures 6(a) and 6(b) reach the predetermined accuracy within T_{max}, but the training times are very different when they reach the predetermined accuracy. Figure 6(a) takes 1178 steps to reach the predetermined accuracy, and Figure 6(b) takes 544 steps to reach the predetermined accuracy.

In the process of network learning, although the initialization of connection weights and thresholds is set randomly, they have a very important impact on the training process of evaluation model, which is embodied in the convergence efficiency and classification accuracy of evaluation model. Under the same conditions, when the initial network weights and thresholds are different, some training processes can converge within the set T_{max} range, while some learning processes cannot reach the predetermined accuracy after learning T_{max} times. In this case, it is necessary to regenerate the network weights and retrain. In the process of practical application, it is often necessary to initialize the network connection weights and thresholds for many times, and compare the training process and test results for many times to obtain the most reasonable evaluation model. Table 3 shows the initial weight training and test results of different networks.

4.4. Comparative Experiment. Firstly, a simple three-layer mutation genetic neural network structure is used to evaluate credit risk without adding genetic algorithm. Then, the weight and threshold of neural network are optimized by the mutation genetic algorithm. Its operation speed shows greater advantages in speed and quality than BP neural network, and achieves higher accuracy. The training results are shown in Figure 7. Comparing the training time with the original data, Table 4 is obtained.
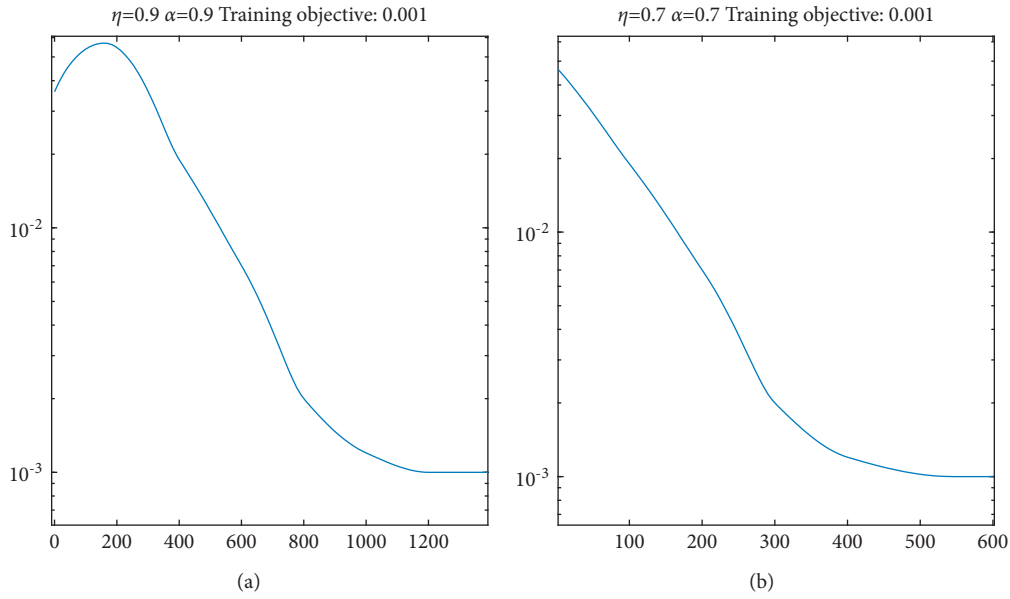
FIGURE 6: Training error function under different η and α. (a) Training times: 1178; (b) training times: 544.

TABLE 3: Model training under different network initial weights.

Mode code	Network weight sequence number	Simulation convergence	Number of classification errors
	Initialization 1	Within 3500 training times, the simulation does not converge	15
0101	Initialization 2	Simulation convergence, training steps 1875	12
	Initialization 3	Simulation convergence, training steps 1168	10
	Initialization 4	Simulation convergence, training steps 1050	11
0102	Initialization 5	Simulation convergence, training steps 978	9
	Initialization 6	Within 3500 training times, the simulation does not converge	11
	Initialization 7	Within 3500 training times, the simulation does not converge	14
0103	Initialization 8	Simulation convergence, training steps 2238	10
	Initialization 9	Within 3500 training times, the simulation does not converge	12
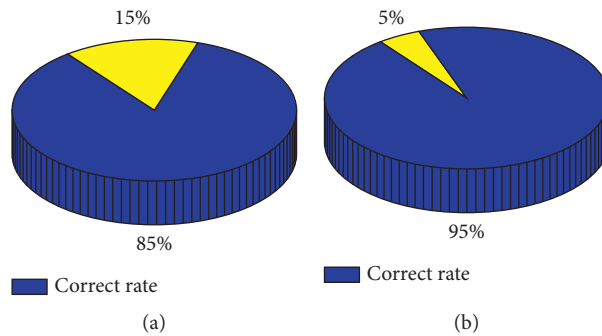
FIGURE 7: Comparison of training accuracy between the two algorithms. (a) BP algorithm; (b) our algorithm.

From its fast operation speed and high accuracy, it can be concluded that the variant genetic neural network has strong advantages in the evaluation of credit risk of commercial banks.

The improved mutation genetic algorithm evaluation model has not only improved the training speed, but also greatly improved the classification accuracy. First select 1000 times to see whether the final training reaches the target error. Then increase or decrease the training times according to the situation. If the training times are the same each time, the degree of training results is very small. It is recommended to increase the number of training times. In the experiment, 2000 times is the best training times. Under the same training steps (set as 2000 steps in this paper), the error

TABLE 4: Comparison of training efficiency between the two algorithms.

	BP algorithm	Our algorithm (s)
Training time	7.230 s	4.836
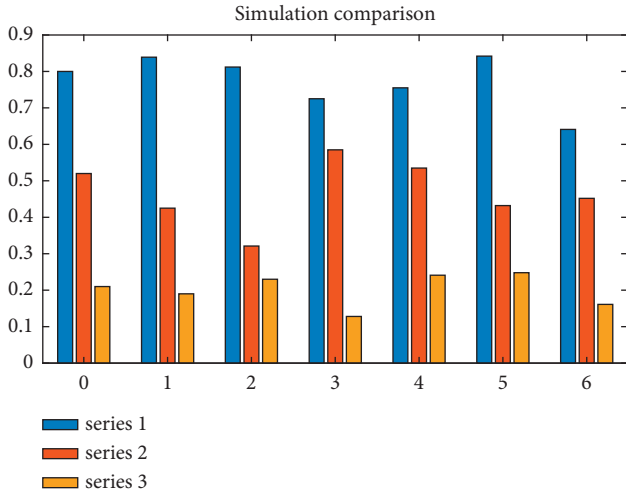
FIGURE 8: Simulation comparison of evaluation models under two algorithms.

accuracy of the basic BP algorithm evaluation model can just reach the accuracy of 0.001, while the error accuracy of the improved mutation genetic algorithm evaluation model can reach 0.0001 at 2000 steps. As shown in Figure 8, the simulation comparison diagram of the evaluation models under the two algorithms is shown, in which series 1 is the expected output, series 2 is the improved algorithm's simulation output, and series 3 is the simulation output of the unmodified algorithm. The number of classification error samples of the evaluation model is 4 by using the unmodified BP algorithm, while the number of classification error samples of the evaluation model is reduced to 2 by using the improved mutation genetic algorithm.

5. Conclusion

The credit risk assessment model established in this paper adopts a three-layer mutation genetic neural network structure. For different assessment models, the assessment model has different network structures, which is embodied in the different number of input layer nodes, hidden layer nodes, and their network connection weights. Taking industrial enterprises as an example, the experimental results show that there are two evaluation modes of credit risk: small industrial enterprises and large- and medium-sized industrial enterprises. Their network structure is the same, $28 \times 18 \times 1$. In the learning algorithm, this paper uses the combination of variable learning rate and momentum term to improve the training efficiency and learning rate of the evaluation model η value 0.8, momentum factor α value 0.9, and error accuracy ε set to 0.001. According to the above network structure and learning parameters, the credit risk

assessment model is established, and the qualitative index standardized data and quantitative index standardized data are used to train the two assessment models, respectively. Due to the randomness of connection weight initialization, this paper initializes and trains each mode for many times and takes the model with the best test accuracy as the final model. Under the same conditions, the training convergence speed of the improved mutation genetic algorithm evaluation model is very different from that of the basic BP algorithm evaluation model. The training speed of the improved BP algorithm evaluation model has been qualitatively improved. The basic BP algorithm evaluation model takes 1178 steps to meet the accuracy requirements, while the mutation genetic algorithm evaluation model takes only 544 steps to meet the progress requirements. It can be seen that the credit risk assessment model based on neural network has high accuracy and can be used as an important tool for loan decision making of commercial banks.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] I. Serwadda, "Impact of credit risk management systems on the financial performance of commercial banks in Uganda," *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 66, no. 6, pp. 1627–1635, 2018.

[2] S. R. Poudel, "Assessment of credit risk in Nepali commercial banks," *Journal of Applied and Advanced Research*, vol. 3, no. 3, pp. 65–72, 2018.

[3] X. Zhou, X. Yang, J. Ma, and K. I.-K. Wang, "Energy efficient smart routing based on link correlation mining for wireless edge computing in IoT," *IEEE Internet of Things Journal*, May 2021.

[4] P. Supraja, V. M. Gayathri, and R. Pitchai, "Optimized neural network for spectrum prediction using genetic algorithm in cognitive radio networks," *Cluster Computing*, vol. 22, no. 1, pp. 157–163, 2019.

[5] T. Rashid, S. M. Abdullah, and S. Abdullah, "A hybrid of artificial bee colony, genetic algorithm, and neural network for diabetic mellitus diagnosing," *ARO-The Scientific Journal of Koya University*, vol. 6, no. 1, pp. 55–64, 2018.

[6] M. N. Kamel Boulos, Z. Lu, P. Guerrero, C. Jennett, and S. Anthony, "From urban planning and emergency training to Pokémon Go: applications of virtual reality GIS (VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health," pp. 1–11, 2017.

[7] Z. H. Wang, D. Y. Gong, X. Li, G. T. Li, and D. H. Zhang, "Prediction of bending force in the hot strip rolling process using artificial neural network and genetic algorithm (ANN-GA)," *International Journal of Advanced Manufacturing Technology*, vol. 93, no. 9, pp. 3325–3338, 2017.

[8] M. Jahanbakht, W. Xiang, L. Hanzo, and M. R. Azghadi, "Internet of Underwater Things and big marine data analytics – a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 904–956, 2021.

[9] Y. Liu, X. Zhu, and J. Yang, "Fault diagnosis of PV array based on optimised BP neural network by improved adaptive genetic algorithm," *Journal of Engineering*, vol. 2017, no. 13, pp. 1427–1431, 2017.

[10] Z. Wan, Y. Dong, Z. Yu, H. Lv, and Z. Lv, "Semi-supervised support vector machine for digital twins based brain image fusion," *Frontiers in Neuroscience*, vol. 15, p. 802, 2021.

[11] S. Shaghaghi, H. Bonakdari, A. Gholami, I. Ebtehaj, and M. Zeinolabedini, "Comparative analysis of GMDH neural network based on genetic algorithm and particle swarm optimization in stable channel design," *Applied Mathematics and Computation*, vol. 313, pp. 271–286, 2017.

[12] S. Xie, Z. Yu, and Z. Lv, "Multi-disease prediction based on deep learning: a survey," *Computer Modeling in Engineering & Sciences*, vol. 128, no. 2, pp. 489–522, 2021.

[13] N. Zhang and J. Zhang, "The application of amended KMV model in measuring credit risk of China listed companies," *Forecasting*, vol. 29, no. 5, pp. 48–52, 2010.

[14] M. Hu, Yi Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers*.

[15] J. Tarigan and Y. Suryana, "Plate recognition using back-propagation neural network and genetic algorithm," *Procedia Computer Science*, vol. 116, pp. 365–372, 2017.

[16] H. Chung and K.-s. Shin, "Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction," *Neural Computing & Applications*, vol. 32, no. 12, pp. 7897–7914, 2020.

[17] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 766–775, 2020.

[18] N. M. Hewahi and E. Abu Hamra, "A hybrid approach based on genetic algorithm and particle swarm optimization to improve neural network classification," *Journal of Information Technology Research*, vol. 10, no. 3, pp. 48–68, 2017.

[19] M. Bahiraei, S. Nazari, H. Moayedi, and H. Safarzadeh, "Using neural network optimized by imperialist competition method and genetic algorithm to predict water productivity of a nanofluid-based solar still equipped with thermoelectric modules," *Powder Technology*, vol. 366, pp. 571–586, 2020.

[20] Z. Yan and Z. Lv, "The influence of immersive virtual reality systems on online social application," *Applied Sciences*, vol. 10, no. 15, p. 5058, 2020.

[21] P. d. S. L. Alexandrino, G. F. Gomes, and S. S. Cunha Jr., "A robust optimization for damage detection using multi-objective genetic algorithm, neural network and fuzzy decision making," *Inverse Problems in Science and Engineering*, vol. 28, no. 1, pp. 21–46, 2020.

[22] J. H. Rhee, S. I. Kim, K. M. Lee, M. K. Kim, and Y. M. Lim, "Optimization of position and number of hotspot detectors using artificial neural network and genetic algorithm to estimate material levels inside a silo," *Sensors*, vol. 21, no. 13, p. 4427, 2021.

[23] M. Elsisi, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "An improved neural network algorithm to efficiently track various trajectories of robot manipulator arms," *IEEE Access*, vol. 9, pp. 11911–11920, 2021.

[24] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.

[25] M. Nazari-Heris, B. Mohammadi-Ivatloo, and A. Haghrah, "Optimal short-term generation scheduling of hydrothermal systems by implementation of real-coded genetic algorithm based on improved Mühlenbein mutation," *Energy*, vol. 128, pp. 77–85, 2017.

Research Article

A Novel Framework for Arabic Dialect Chatbot Using Machine Learning

Nadrh Abdullah Alhassan,[1] Abdulaziz Saad Albarrak ⓘ,[1] Surbhi Bhatia ⓘ,[1] and Parul Agarwal ⓘ[2]

[1]*Department of Information Systems, College of Computer Science and Information Technology, King Faisal University, Alhasa, Saudi Arabia*
[2]*Department of Computer Science and Engineering, Jamia Hamdard, New Delhi, India*

Correspondence should be addressed to Surbhi Bhatia; sbhatia@kfu.edu.sa

With the advent of artificial intelligence and proliferation in the demand for an online dialogue system, the popularity of chatbots is growing on various industrial platforms. Their applications are getting widely noticed with intelligent tools as they are able to mimic human behavior in natural languages. Chatbots have been proven successful for many languages, such as English, Spanish, and French, over the years in varied fields like entertainment, medicine, education, and commerce. However, Arabic chatbots are challenging and are scarce, especially in the maintenance domain. Therefore, this research proposes a novel framework for an Arabic troubleshooting chatbot aiming at diagnosing and solving technical issues. The framework addresses the difficulty of using the Arabic language and the shortage of Arabic chatbot content. This research presents a realistic implementation of creating an Arabic corpus for the chatbot using the developed framework. The corpus is developed by extracting IT problems/solutions from multiple domains and reliable sources. The implementation is carried forward towards solving specific technical solutions from customer support websites taken from different well-known organizations such as Samsung, HP, and Microsoft. The claims are proved by evaluating and conducting experiments on the dataset by comparing with the previous researches done in this field using different metrics. Further, the validations are well presented by the proposed system that outperforms the previously developed different types of chatbots in terms of several parameters such as accuracy, response time, dataset data, and solutions given as per the user input.

1. Introduction

Every generation of the computer device is becoming more complex than the last one, and the troubleshooting task becomes more difficult for the end-user [1]. Computer problem detection is a complicated process that demands a high level of knowledge and skills. Troubleshooting is the process of locating the cause of a problem in a system and resolving it [2]. On average, computer repair technicians charge globally $60 per hour. Besides, the hourly rate can range from $40 to $90 per hour [3]. Some of these repairs can be addressed by providing assisting through a chatbot to the user. Furthermore, one of the remarkable artificial

intelligence applications that have proven its efficiency recently is a chatbot. The chatbot is a simulating program that analyzes and processes human conversation, either written or spoken. It interacts with the end-users via digital devices as if users are communicating with a real person [4]. It provides 24-hour availability, instant answers, endless patience, personalization, and customized dialog to the end-user. Additionally, it extends to benefit various companies in this matter by saving the cost of hiring manpower, achieving customer satisfaction, and reaching new customers [5]. Chatbot relies heavily on Natural Language Processing (NLP), commonly used to make human spoken language understandable for computer machines [6]. The goal of NLP

is to take the unstructured input and produce a structured representation of the text that contains understandable language for textual chatbot conversation [7]. Also, Natural Language Processing is known as a model trained which belongs to an Arabic chatbot's textual input and output [8]. The essential part of building chatbots is the conversational interfaces. Since the chatbot responds in terms of text to the user query, NLP yields an adequate textual conversation for the chatbot [9]. In this paper, we rely on NLP to process Arabic language by using the presented framework. The research has identified and minimized the gaps existing in the field of dialects in Arabic language and the advancements in the chatbots by comparing the performance of the proposed system with the previously developed researches.

The rest of this paper is organized as follows: next, we present motivation and contributions and a tiny backdrop. We discuss the challenges for developing Arabic chatbots. We then present comparison with existing systems and describe our framework. After that, we illustrate chatbot mechanism used in this paper including experimental result and a preliminary user evaluation.

1.1. Motivation and Contributions. The motivation behind this work has portrayed the fact that Arabic chatbot does not reach the required limit in terms of response speed, dataset type, and how the chatbot handles user input. The outcome of this paper will lie in the implementation based on industrial applicability in building an Arabic Chabot that is capable of assisting end-users to solve troubleshooting problems in Arabic. This framework incurs repair costs by an expert troubleshooting system. It will aid to diagnose IT troubleshooting and solve technical issues within a few seconds. The contributions of this research work are given as follows:

 (i) To study the background on previously developed chatbots and the challenges of using Arabic chatbot in multiple domains;

 (ii) to propose the novel IT troubleshooting chatbot framework for Arabic language;

 (iii) to create the dataset from various sources including multiple companies such as Microsoft, HP, Samsung, Huawei, and different websites that address common IT problems/solutions and apply preprocessing tasks for further uses;

 (iv) to compare the proposed chatbot with previous researches on varied dialogues, writing the parameters;

 (v) to evaluate the proposed chatbot by comparing it with other chatbots on different functional evaluation metrics;

 (vi) to develop a service-based chatbot to provide idiomatic solutions for common IT problems that can be considered for practical industrial applications.

2. Background

The revolution of information technology during our era predetermines that all governmental and private sectors need rapid progress of digital transformation and mounting development. Our lives are not devoid of using technologies, especially those that involve artificial intelligence techniques. Technology has gathered the world societies in a common cyber environment, where every technology that emerges in any place spreads to all applicable societies.

Artificial intelligence relies on data as input; then it is processed via specific approach to generate the demand output like-human thinking. Even though, there is a concern whether AI application will rescind the need of human thinking gradually. In fact, AI works as extension of human brain to solve complex problems and process huge amount of data at the same time. One of AI techniques used with different language by machine learning is the acknowledgment of city names. The promising study application areas in the realm of postal automation are the recognition of handwritten city names. For recognition, use a nonsegmentation method (Holistic approach). The role of the convolutional neural network (CNN), which is one of the deep learning techniques, is deconstructed in this paper [8], detecting hand signs, by anticipating the next word or recommending the most relevant word, and then generating the word that deaf persons communicate with people using sign language.

2.1. Arabic Chatbot. Before time, chatbot was described as a text-interaction between human and machine learning by simulating an online conversation (chat) with a user in natural language. Chatbot was first introduced in 1996. Precisely, ELIZA was the first chatbot created in MIT, which arguably came close to imitating a human reply. ELIZA was given an input sentence and it identifies keywords and patterns to match those keywords against a set of preprogrammed rules to generate appropriate responses [10]. Late in 1971, Kenneth Colby at Stanford created PARRY, a preprogrammed chatbot act-like schizophrenia diagnosed patient and was able to express fears, anxiety, and beliefs. By the time, chatbot's discoveries increased in 1995 by Richard Wallace who created A.L.I.C.E chatbot, using English conversation patterns in AIML files such as (<category>...<pattern></pattern>....</category>). AIML is a subordinate of web extensible mark-up language XML. After that, during the millennium, many forms of chatbot were invented. To get a sense, all millennium chatbot inventions refer now to modern chatbots in the AI industry.

The Arabic language is considered as one of the Afroasiatic languages. Due to built-in alphabetical structure, it is a notoriously difficult language in categories of programming text processing, because of reading and writing direction, from right to left, which is the opposite direction of English language. To further complicate things, the form of Arabic letter is totally different than English language and vowels are omitted from written Arabic. Globally, there are 422 million Arabic speakers around the world [11]. According to a British council report, Arabic language took

the second place for each native speaker who lived in northern Africa such as Algeria and Morocco and western Asia such as Georgia and Azerbaijan and took the first place at Arabian Peninsula such as Saudi Arabia and United Arab Emirates. Moreover, it took the 4th place as a useful language in the trade market in the UK and Arabic countries as well [12].

In this respect, significant work has been done on chatbots. However, the Arabic language is rarely used in chatbot especially in the IT sector. Additionally, few numbers of Arabic chatbots applications/websites are available for the end-user. For instance, in educational domains, Nabiha, one of Arabic chatbots, is concerned about helping college students using informal Arabic via automatic conversation with student inquiries [13]. On the other hand, BOTTA is a fun Arabic conversational chatbot aiming at using an Egyptian accent to have fun chat with users [14]. Also, Quran chatbot is an Arabic one answering users' inquiries religiously, by generating extracted replies from the Holy Quran [15]. Also, Tafsir Al-Ahlam is an Arabic robot specialist interpretation of dreams using a smart search engine to generate output from a local database of a wide range of interpretations taken from the books attributed to Ibn Sirin and Nabulsi son of Shaheen and other accredited authors in the Islamic heritage books, following the origins and rules of interpretation of dreams from the Quran and the Hadith [16]. In medical domain, NALA is an application that provides medical consulting under the supervision of the Ministry of Health [17].

Customers may use service-oriented chatbot to find information on big, complex domains that are difficult to navigate. This is recognized as "service-based chatbot;" the main reasonability for this chatbot is to provide a service, contrary to other chatbots acting as only an entertainer. Many users find it difficult to find the details that they need from website search engine results due to the site's abundance of data. During a conversation with the user, the service-oriented chatbot serves as an automated customer service agent, providing natural language responses and more targeted information. This virtual agent is also programmed to assist with general IT-related questions [18].

2.2. Chatbot Framework.
A framework is a term commonly used in IT software development which is referred to as a structure consisting of many phases that work together as a foundation for software developers to build programs for a specific platform [19]. As it follows by bot framework, a set of processes and requirements guide the developer while he/she is programming the bot as well as solving any potential problems met by the developer at the same phase. Bot framework provides tools, services, and skills needed which facilitates the developer's job [20].

Most frameworks share certain fundamentals phases, but they differ according to the purpose of creating the framework itself, based on the most proposed framework that is currently going on in the field of chatbot development. A five functional phases are listed as follows. Automatic speech recognition (ASR) is a technology that provides a voice identifier with a computer interface through the human voice that allows a human being to interact with a computer in such a way that it seems very close to the real human. Natural language understanding (NLU) is highly connected to machine reading comprehension. The process of selecting parts of sentences and analyzing the meaning is tedious because the machine needs to determine the correct syntactic structure and semantics of the language used. Dialogue management is mainly responsible for coordinating the various components of a chatbot. Natural language generation is mainly relating to the process of retrieving and producing the answer. Lastly, there is text to speech. The answer is ready to be retrieved for the end-user [21]. The following is a brief description of some frameworks used for chatbots in Arabic.

2.3. Comparison with Existing Systems.
The first framework is the framework of an intelligent Arabic chatbot for teaching Islamic history. The main goal of this framework is to help and guide designers for the efficient use of chatbots and simulations for teaching Islamic history. It consists of three phases explained as follows.

2.3.1. Interaction.
This phase includes the user input such as query, and it is stored immediately in the Heroku cloud server. This framework adopts a cloud computing server to avoid standard dedicated server's problems and costs.

2.3.2. The Processing Phase.
This phase includes the use and implementation of many programming languages for handling, storing, processing, searching, and retrieving the data. After that, determining the size of storage might be needed since the content of Islam history is too large to be handled on smartphone's capacity. So, the framework presents the MongoDB technique to be used in this phase.

2.3.3. Learning Phases.
This phase enables the bot to be recordable and learnable, by sorting user input/query at bot memory. The program manager will update the bot's memory by matching words or queries with the appropriate answer. Finally, the last process is using some advanced programs such as neurolinguistic programming (NLU), and it can be implemented manually by human review or automatically by NLU. The current NLU program cannot support Arabic language but it can be replaced with any suitable text processing techniques [22].

The second framework is Facebook Arabic chatbot based on deep learning using RASA framework aimed at answering student inquiries at the University of Islam in Indonesia via Facebook message. College students often need immediate information like asking for something to help desk, especially during this COVID-19 pandemic, due to Facebook popularity in Indonesia. There were 166.500.000 Facebook users in Indonesia in August 2020, which accounted for 60.8% of its entire population. Thus, RASA framework has been developed to match with Arabic content [23].

To construct an Arabic chatbot, most of the existing research relies on third-party framework platforms likewise PANDORABOT platform. As a result, some difficulties in dealing with Arabic letters have arisen as well as HTML tags, database scope, response, and input processing. The following framework will overcome these gaps by using programming Python groundwork. In terms of Arabic content, the framework uses AI to improve NLP and machine learning techniques; see Table 1.

As shown in Table 1, most of the Arabic chatbot applications used premade platform like Pandorabots API. This kind of platform has subscribed package followed by yearly/monthly fees. The developer is in charge of renewing the subscription fee. Eventually, the developer gets exhausted with unnecessary additional costs and limited features, gaining complexity at run time. This conceded on the top list of solutions provided by the proposed framework. Also, code-based application frameworks are rarely supported with Arabic content. This research focuses on solving the problems of complexity and user-friendly programming support with effective applications.

3. The Proposed Methodology

This research introduces a novel IT chatbot framework for troubleshooting supporting Arabic language. The framework comprises four phases: GUI, Arabic text processing, AI services, and database. Figure 1 demonstrates the components and the processes involved in the framework design.

The chatbot starts by receiving the input as "text format" from the end-user. The bot channels transfer the question/ inquiry regarding any potential IT troubleshooting as a query to the Arabic text processing phase. At this stage, the framework examines the match corpus from the existing knowledge in the database phase and extracts the answer as result to the end-user. Along with that, get the benefit from AI services to enhance the learning process in the database phase for nonexisting knowledge and improve generally text process in Arabic text processing phase. Text processing phase not only provides NLP, but also provides named entity recognition (NER) by labelling named "real-world" objects, like persons, companies, or locations. Entity linking (EL) removes the uncertainty of meaning from an ambiguous sentence, phrase, or other linguistic units. Text processing is useful for more than just decoding and analyzing text back to its origins. Rather, it goes above and above to generate data that aids in the usage of chatbot measures. As it is a real number, it is very easy to do statistical and arithmetical calculations towards NLP, making it suitable for use in statistical ML models and finally using bot channels to retrieve the solution in GUI format through a friendly interface. Moreover, one of the goals intends to contribute to the NLP uses of the Arabic language. The proposed framework endeavor deeply presents how to process Arabic text in the IT domain.

Moreover, every chatbot has its dataset or in other terms corpus. The Quran chatbot is one of the Arabic chatbots that adapted the Java program to produce an Arabic AIML dialogue corpus [15]. BOTTA is another Arabic application using AIML files through the Pandora's platform to generate a dialogue corpus [14]. However, there is no Arabic chatbot that has been adopted in Python programming language and generates the corpus via Python environment. So, we built a service-based chatbot to provide idiomatic solutions for common IT problems using python. Besides, build a knowledge base of IT troubleshooting, which can be reused by other researchers.

3.1. Chatbot Builder. The work has been implemented on a personal laptop with specific features. To deal with the capabilities of the chatbot system, CPU, which is a Dual-Core Intel Core i7 with 8 GB RAM, is used. The processing speed of the CPU is essential in the programming and data mining phases. The mean programming language used is Python and any related library/package is used to run the chatbot mechanism.

3.2. Dataset. The incredible amount of data on the Internet is a rich resource for any chatbot application. Web scraping is the process used in this project to transform unstructured data into a specific classification, thus to extract data using web scraping with Python including nltk library to reprocessing Arabic content. And use an open-source project corpus for the Saudi dialect called MADRA PROJECT by data analysts from Jordan. A corpus is a selection of different input statements and responses for the bot to practice with. This project allows Arabic developers to use more than 26 Arabic dialects [24]. But, here in this project, we build our corpus called chattest.text, with the benefit of merging MADRA's corpus and our corpus.

3.3. Arabic Text Handling. By using online 50 contents as database source, whether structured as csv or unstructured as text/yml, some problems occur. Thus, we apply text preprocessing with nltk: tokenization, by splitting the Arabic text into smaller pieces or "tokens." **Normalization** aims to put all Arabic content in one level. Noise removal cleans up the text from extra white spaces.

Tables 2 and 3 list the details on the dataset and websites used for extracting the data. In total, 50 texts have been retrieved with nested content, with difficulty in forming data according to the classification prepared previously in the research and some python's library/package, not supporting Arabic language. But by using core web sm for the latest version and scan of the HTML tag of required content, use the HTML attribute to specify the content.

The result of data after applying the classification process is shown in Figures 2 and 3.

3.4. Chatbot Mechanism. After generating a dataset that is suitable for the IT problem/solution domain, the implementations were done using Python programming language. Python is more suited to make changes to an existing legacy system and offers different methods/libraries such as NumPy, pandas, PyBrain, and SciPy that help expedite AI development. For example, you can leverage proven

TABLE 1: Comparison of a few of the relevant works pertaining to the Arabic chatbots.

Ref no.	Motivation	Datasets used	Obtained accuracy	Limitations
[13]	This article introduces Nabiha, a new Arabic dialect chatbot dedicated to assisting college students	The datasets collected are 248 inputs/outputs from the KSU IT students' accounts in Askme.com. The methods used in this work are premade platform which is called PANDORABOT platform	The result of this paper shows that 15.38% of the total answer is not accurate, 53.85% is somewhat accurate, and 30.77% is very precise	The dataset should be expanded; we need to address the issue of HTML tags as well as Twitter's text area constraints
[14]	BOTTA's goal is to create a conversational environment and connect with as many Arab users as possible. She's the first chatbot to speak in an Arabic dialect, which helps her achieve her goal of amusing people who are used to conversing in the language	BOTTA is using AIML and launched it on the PANDORABOT platform	BOTTA's pattern matching will be able to correct 85.1 percent of the spelling errors observed in spontaneous Arabic typing	Using corpus-based machine learning approaches, BOTTA's pattern matching has to be improved. Further development will involve morphological analysis of the input and experiments with lemma-based pattern matching using existing tools for Egyptian Arabic processing
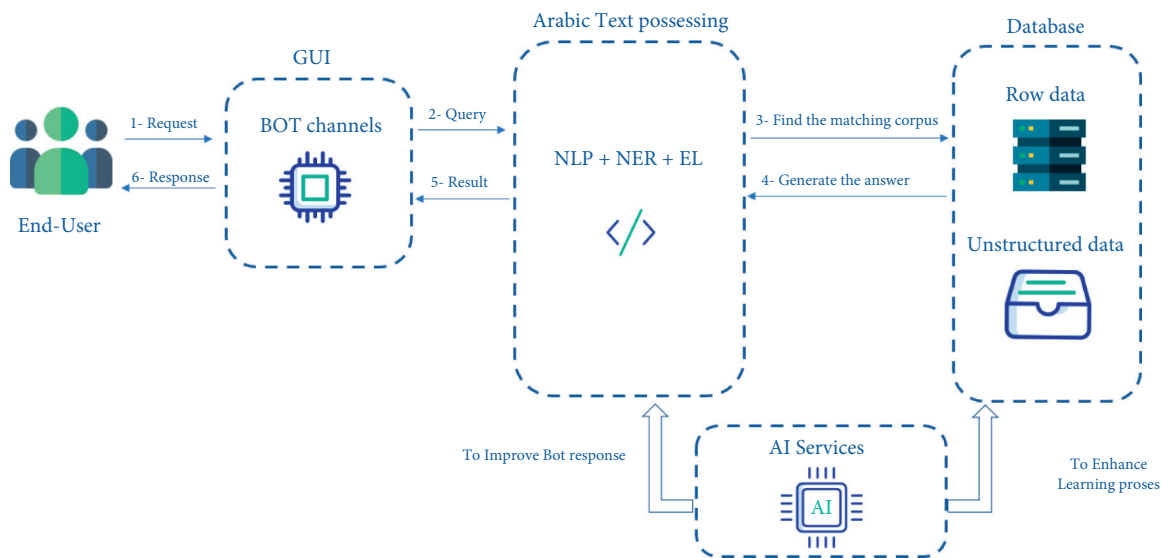
FIGURE 1: System proposal framework of devolving an Arabic troubleshooting chatbot to diagnose and solve technical issues.

libraries like scikit-learn for ML and use regularly updated libraries like Apache MXNet, PyTorch, and TensorFlow for DL and bots' projects that support text possessing in Arabic field.

Thus, we create a greeting message for the chatbot and then create (AI keyword matching) for the chatbot's corpus. We use (AI) intelligent search through the corpus using spacy and Python environment. Keyword matching works on using a keyword that appears in the query. And it identically matches any word in a chatbot corpus. In a basic retrieval system, keyword matching cannot be functional without having the full sentence or query to retrieve the relevant answer. But, by using AI keyword matching, the chatbot does not analyze the whole user input but focuses on searching words on phrases defined in the user says [25]. The best way to explain the behaviour of AI keywords is to use a

realistic example as shown in Table 4. In case the chatbot cannot find the entry that matches a keyword, it will return: "I'm sorry! I do not understand you."

4. Experimental Results

The proposed framework is evaluated in terms of effectiveness. The comparison of the chatbot developed is compared with other chatbot's platforms using different parameters including time of response, type of dataset, and dealing with user input. The two versions of the chatbot are developed, one by using the proposed framework and the other by using an external framework/platform which is Pandorabot [26]. Both chatbots were fed by the same datasets. To measure the variation between the two bots, different experiments have been performed. Table 5 shows

TABLE 2: Some of the open-source websites with the common companies for IT troubleshooting.

Name of company	Name of the website	Does it provide Arabic content?
Microsoft	https://support.microsoft.com/ar-sa/windows/windows-update-%D8%A7%D9%84%D8%A3%D8%B3%D8%A6%D9%84%D8%A9-%D8%A7%D9%84%D9%85%D8%AA%D8%AF%D8%A7%D9%88%D9%84%D8%A9-8a903416-6f45-0718-f5c7-375e92dddeb2	Yes
HP	https://support.hp.com/emea_middle_east-ar	Yes
Samsung	https://www.samsung.com/sa/support/mobile-devices/how-do-i-solve-the-problem/	Yes
Apple	https://support.apple.com/ar-sa	Yes
Huawei	https://consumer.huawei.com/sa/support/	Yes

TABLE 3: Some open-source websites used to gather data for IT troubleshooting.

Name of the website	Does it provide Arabic content?	Does it need translation?
https://trendsarab.com/post/4385/%D9%85%D8%B4%D8%A7%D9%83%D9%84-%D8%AA%D9%82%D9%86%D9%8A%D8%A9-%D8%B4%D8%A7%D8%A6%D8%B9%D8%A9	Yes	No
https://www.arageek.com/tech/how-to-deal-with-windows-problems	Yes	No
https://www.youm7.com/story/2019/6/26/8-%D9	Yes	No

FIGURE 2: Final form of data scraped after running the code categorized.

FIGURE 3: Final form of data scraped after running the code categorized.

TABLE 4: Presentation of an example of an Arabic dialogue implemented by the chatbot.

Dialogue in Arabic

"مساعد الذكي: آسمي المساعد الذكي,أذكر لي مشكلتك؟ وسأحاول جاهدا لحلها"

"مستخدم: أريد إزالة تحديث في جهازي"

Here, the user uses only 2 words that match with chatbot corpus: "إزاله" and "تحديث" and still the chatbot diagnoses the problem by saying "إزالة تحديث مثبت في ويندز"

مساعد الذكي: إزالة تحديث مثبت في ويندز:

And retrieve the most relevant answer:

"وصي بإزالة أي من التحديثات التي تم تثبيتها ومع ذلك، إذا كان من الضروريّ إزالة أحد التحديثات، فيمكنك القيام من ال بذلك في محفوظات التحديثات"

Translation of the dialogue

Chatbot: hi, my name is "the smart assistant" tell me your technical problem? And I will try to help you

User: I want to remove an update on my device

Chatbot: we do not recommend removing any installed updates. However, if it is necessary to remove an update, you can do it in the update history."

TABLE 5: Comparison of the chatbots (our chatbot and Pandorabot).

Comparison criteria	Chatbot produces by using our proposal framework	Chatbot produces by using Pandorabot platform
Type of dataset	- Use unstructured Arabic data (text/yaml). And there is no need to define each question/answer because our chatbot can search through all corpus within seconds to retrieve the data.	- Use AMIL file required to define each question as `<pattern></pattern>` and the answer as `<template></template>` for each dialog. - The question must be in capital letter (English case). - The data must not have any kind of punctuations because the platform will ignore it immediately and will not deploy it.
Time of response	The answer was retrieved immediately.	The answer took around 30 seconds to retrieve, and it could be "apologize message."
Dealing with user input	- The user can insert any Arabic input by using the formal and informal Arabic language.	- The user can insert only the exact question already fixed in the AIML file. Except that, the chatbot will ignore the question and retrieve "apologize message."
The experimentation	We insert the same dataset for both bots. And we use the same question format. The question used is "نسقمة يت مرور جوالي الايفون" and the answer should be "نسقمة كلمة تغيير > والأمن السر كلمة [اسمك] على انقر<الإعدادات>في الايفون" اي دي ابل" مرور كلمة يت السر" which contained diagnosis of the problem followed by the solution.	- If the question is in informal Arabic language, the bot retrieves "apologize message." "نسقمة الايفون جوالي مرور كلمة يت" Input: Output: the bot retrieves error "apologize message." - If the question is in an exact AIML file premade, the bot retrieves the correct answer. Input: نسقمة في الايفون" اي دي ابل" مرور كلمة يت Output: the bot retrieves the correct answer.
The result	- If the question is in informal Arabic language, the bot retrieves the correct answer. Input: "جوالي الايفون سرورم كلمة يت" Output: the bot retrieves the correct answer.	
Finding	- Our chatbot was capable of having any type of Arabic data without having strict inputs to a specific form. The bot uses AI keyword matching to analyze user input and match it with the most relevant problem/solution.	- Pandorasbot working as simple chatbot was capable of matching a text string and offering an answer only when the exact sentence match is found.

the type of dataset used, time of response, and how to process the inputs and then illustrate the question chosen to test both chatbots. After that, there are the result and findings.

5. Evaluation

Goal-oriented chatbots are a form of currently popular chatbot whose main function is to assist users with a specific set of tasks, such as "how can I unlock my Samsung (mobile) device?" Some publications with goal-oriented chatbots conducted evaluations that were related to the chatbot's objectives. The metrics include the number of successful conversations ended by the system, percentage of the problem solved, average dialogue length, and average of user utterance length [27].

The participants comprised 6 participants from the different educated background like engineering, linguistic, and finance background. Two were females, 4 were males, 3 were aged between 25 and 34, and 2 were aged between 35 and 44. All participants who participated were not having any IT background.

Sample of chosen dialog/ participant to be tested on the Functional evaluation phase

The interface of the chatbot dialog:

Dialog 1
Name of task: Assume that you need help Installing the HP printer by using a wired
network,(Kindly use your own Arabic language).

Part 1:

   - Personal information of the participant:

Name: Albandary abdullah
Gender: F
Age: 36
Level of education: Master's degree.
Education/work field: Linguistics.

Part 2:

   - Implement the task by the participant:

FIGURE 4: Goal-oriented chatbots evaluation for "المساعد الذكي" chatbot.

5.1. Goal-Oriented Chatbots Evaluation for "المساعد الذكي" Chatbot. We have presented the key findings of designing an Arabic chatbot unescorted by any third-party platforms, based on our proposed framework. The components of our framework are crafted thoughtfully; each phase can add/upgrade your bot mechanism into AI chatbot using the Arabic language.

Most of the current Arabic chatbots use the basic dialog (Q/A), third-party platforms, and complex AIML files [28]. However, we found that our framework directs the developer to get leverage of AI open-source package. In this project, we adopt a Python library to preprocess over 50 Arabic common IT problems/solutions and trained the bot

TABLE 6: General analytics of goal-oriented evaluation part 1.

Dialogs	Problem solved (yes = 1, no = 0)	Problem solved (first reply = 1/second reply = 0,1/more than 2 replies = 0)	Answer relevant (yes = 1, no = 0)
D1	1	0	1
D2	1	0.1	1
D3	0	0	0
D4	1	1	1
D5	1	1	1
D6	0	0	0

PROBLEM SOLVED BY CHATBOT
IN PERCENTAGE

33%

67%
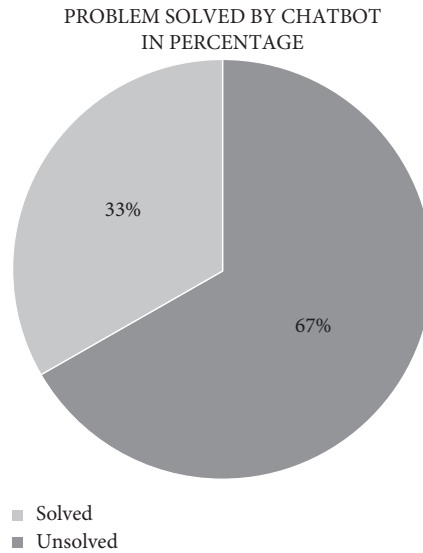
■ Solved
■ Unsolved

FIGURE 5: Solved/unsolved cases presented in pie chart.

ANSWER RELEVENT BY CHATBOT
IN PERCENTAGE

47%

53%

■ Relevent
■ Non-Relevent
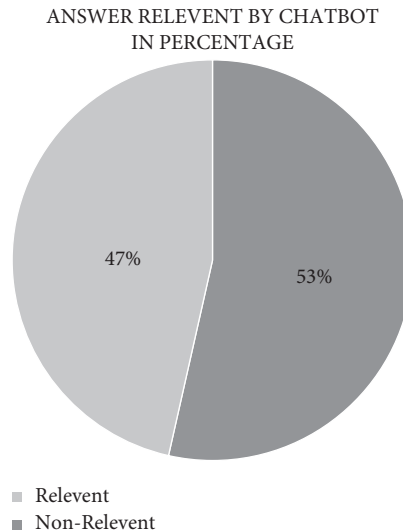
FIGURE 6: Relevant/nonrelevant chatbot's answer presented in pie chart.

to diagnose the problem (input), then generating the solution (output) as shown in Figure 4.

5.2. Performance Measures. The performance of our chatbot has been tested through experiments performed between chatbots, developed from zero starting point, and the other has been tested by using an external platform which is Pandorabot, Arabic troubleshooting chatbot [29, 30]. Both were suited to have any type of Arabic data (formal/informal) without having strict inputs to a specific form. The analysis is shown in Table 6.

TABLE 7: General analytics of goal-oriented evaluation part 2.

Dialogs	Dialogue length (in sentence/line)	User utterance length (in sentence/line)	Number of times in which the participant paraphrases the question
D1	16 sentences	5 sentences	3 times
D2	13 sentences	4 sentences	1 time
D3	17 sentences	8 sentences	4 times
D4	11 sentences	4 sentences	0 times
D5	7 sentences	3 sentences	0 times
D6	14 sentences	7 sentences	3 times
The average	13.17	5.17	1.8

The bot uses AI keyword matching to analyze user input and match it with the most relevant problem/solution. AI keyword matching is a Python package deal [31, 32] with Arabic letters intelligently, needlessly to afford subscription fees to third-party platforms for creating an Arabic chatbot. Figures 5 and 6 show the solved cases in terms of evaluating the performance of the proposed chatbot.

Another testing method of "مساعد الذكي" chatbot is using goal-oriented chatbots evaluation. The result shows that 67% of the total problems are solved successfully and the average dialogue length is around 13 sentences per/conversation. Moreover, the average user utterance length is 5 sentences per/conversation. People usually speak over 8–10 sentences in one minute. Thus, chatbot's conversation does not exceed one minute from the end-user to receive useful information as shown in Table 7. The performance has been evaluated using metrics [33].

6. Conclusions

In this paper, an introductive framework for devolving Arabic troubleshooting chatbot to diagnose and solve technical issues is developed. A framework that supports Arabic text for better machine's understanding is implemented. The framework takes advantage of AI to enhance NLP and machine learning techniques in terms of Arabic content. This chatbot is a service-based chatbot. Thus, 10 websites were scanned to gain 50 problems and their reliable solutions. Chatbot evaluation shows that after comparing two previous built Arabic chatbots on the same dataset the proposed troubleshooting chatbot is capable of diagnosing and answering users' input more effectively. It is more convenient for the users to use their own accent (informal Arabic language) and receive the required answer.

For the evaluation techniques, this research classified them into two main categories, stated as comparison of the chatbot produced by following the proposed framework and others on chatbots platform. The evolution of the chatbot is done by the real users using functional evaluation matrices. Despite the findings of the research, there are some limitations because the majority of Arabic chatbot application has limited type of question answering (QA) corpus via NLP. To make interaction with the Arabic language easier, libraries for programming languages must be developed.

This knowledge is an important basis for future Arabic chatbot development. The framework is developed using the latest released package of Python and Arabic language, by leveraging a variety of AI including keyword matching techniques. However, the Arabic language branched out into special grammar, character, and forms that may be difficult for the Python language to identify in this research. The findings can be used as a guide for potential chatbot developers to choose appropriate evaluation methods for their chatbots based on their design, domain, aim, and evaluation intent. Also, it is possible to add voice bots by using voice-to-text recognition after incorporating the required technology. Furthermore, the outcomes of chatbot evaluations could be compared through various chatbots that were evaluated independently.

Data Availability

This paper is the result of an ongoing project; the dataset is developed and cannot be released at this stage.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

References

[1] F. Bavera and E. Bonelli, "Justification logic and history based computation," in *International Colloquium on Theoretical Aspects of Computing*, pp. 337–351, Springer, Berlin, Heidelberg, 2010.

[2] C. Thorne, "Chatbots for troubleshooting: a survey," *Language And Linguistics Compass*, vol. 11, no. 10, Article ID 12260, 2017.

[3] Thumbtack, "Computer Repair Prices | Mac, PC and Laptop Repair Costs," *Thumbtack*, 2020, https://www.thumbtack.com/p/computer-repair-prices.

[4] M. Dahiya, "A tool of conversation: chatbot," *International Journal of Computer Science Engineering*, vol. 5, no. 5, pp. 158–161, May 2017.

[5] C. Dilmegani, "Top 12 Benefits of Chatbots in 2020: The Ultimate Guide," *Applied Artificial Intellignece*, vol. 29, 20120.

[6] J. Thanaki, "Python Natural Language Processing," *Packet Publishing Ltd*, 2017.

[7] J. Cahn, *Chatbot: Architecture, Design, & development*, Ph.D. University of Pennsylvania, PA, USA, 2017.

[8] S. Sharma, S. Gupta, D. Gupta, S. Juneja, and G. Singal, "Recognition of Gurmukhi Handwritten City Names Using Deep Learning and Cloud Computing," *Scientific Programming*, vol. 2022, Article ID 5945117, 16 pages, 2022.

[9] S. Juneja, A. Juneja, G. Dhiman, S. Jain, A. Dhankhar, and S. Kautish, "computer Vision-Enabled character recognition of hand Gestures for patients with hearing and speaking disability," *Mobile Information Systems*, vol. 2021, Article ID 4912486, 10 pages, 2021.

[10] P. Bii, "Chatbot technology: a possible means of unlocking student potential to learn how to learn," *Educational Research*, vol. 4, no. 2, pp. 218–221, 2013.

[11] V. Vashisht and P. Dharia, "Integrating chatbot application with qlik sense business intelligence (BI) tool using natural language processing (NLP)," in *Micro-Electronics and Telecommunication Engineering*, pp. 683–692, Springer, Singapore, 2020.

[12] S. AlHumoud, A. Al Wazrah, and W. Aldamegh, "Arabic chatbots: a survey," *International Journal of Computer Scinece Applications*, vol. 9, no. 8, pp. 535–541, 2018.

[13] Z. Safi, A. Abd-Alrazaq, M. Khalifa, and M. Househ, "Technical aspects of developing chatbots for medical applications: scoping review," *Journal of Medical Internet Research*, vol. 22, no. 12, Article ID 19137, 2020.

[14] A. Shawar and E. S. Atwell, "An Arabic chatbot giving answers from the Qur'an," in *Proceedings of TALN04: XI Conference sur le Traitement Automatique des Langues Naturelles*, vol. 2, pp. 197–202, Fez, Morocco, April 2004.

[15] D. Al-Ghadhban and N. Al-Twairesh, "Nabiha: an Arabic dialect chatbot," *International Journal of Advance Computer Science Applications*, vol. 11, no. 3, 2020.

[16] D. Abu Ali and N. Habash, "Botta: an Arabic dialect chatbot," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pp. 208–212, Osaka, Japan, December 2016.

[17] N. A. Al-Madi, K. A. Maria, M. A. Al-Madi, M. A. Alia, and E. A. Maria, "An intelligent Arabic chatbot system proposed framework," in *Proceedings of the 2021 International Conference on Information Technology (ICIT)*, pp. 592–597, Amman, Jordan, June 2021.

[18] "شات بوت لتفسير الاحلام – Apps on Google Play," *Tafseer Alahlam Chatbot*, 2018, https://play.google.com/store/apps/details?id=com.tfserahlaam&hl=en&gl=US.

[19] Nala Company, "الان - طبيب مع تحدّث 'Speak with Your Doctor'," 2020, https://play.google.com/store/apps/details?id=com.tfserahlaam&hl=en&gl=US.

[20] P. Smutny and P. Schreiberova, "Chatbots for learning: a review of educational chatbots for the Facebook Messenger," *Computers & Education*, vol. 151, Article ID 103879, 2020.

[21] E. S. AL-Hagbani and M. B. Khan, "Support of existing chatbot development framework for Arabic language: a brief survey," in *Proceedings of the 5th International Symposium on Data Mining Applications*, pp. 26–35, Springer, Riyadh, Saudi Arabia, March 2018.

[22] H. Nukala, *A Chatbot Framework for Yioop*, Ph.D. Master of Science, San Jose State University, San Jose, CA USA, 2017.

[23] S. Chandel, Y. Yuying, G. Yujie, A. Razaque, and G. Yang, "Chatbot: efficient and Utility-based platform," *Advances in Intelligent Systems and Computing*, vol. 1, pp. 109–122, 2019.

[24] I. Aijaz and P. Agarwal, "A study on time series forecasting using hybridization of time series models and neural networks," *Recent Advances in Computer Science and Communications*, vol. 13, no. 5, pp. 827–832, 2020.

[25] E. Atwell, C. Brierley, K. M. Dukes, and A. B. Sharaf, "An Artificial Intelligence approach to Arabic and Islamic content on the internet," in *Proceedings of NITS 3rd National Information Technology Symposium*, pp. 1–8, Riyadh, Saudi Arabia, May 2011.

[26] N. Habash, "The MADAR Project," 2007.

[27] M. Honnibal, "Rule-based Matching ● spaCy Usage Documentation," 2021.

[28] N. S. Raj and B. R. Suteja, "Implementasi AIML pada Pandorabot untuk studi kasus fakultas teknologi informasi," *Jurnal STRATEGI-Jurnal Maranatha*, vol. 1, no. 1, pp. 13–22, 2019.

[29] W. Maroengsit, T. Piyakulpinyo, K. Phonyiam, S. Pongnumkul, P. Chaovalit, and T Theeramunkong, "A survey on evaluation methods for chatbots," in *Proceedings of the 2019 7th International Conference on Information and Education Technology - ICIET 2019*, pp. 111–119, Aizu-Wakamatsu, Japan, March 2019.

[30] S. Mohamad Suhaili, N. Salim, and M. N. Jambli, "Service Chatbots: A Systematic Review," *Expert Systems with Applications*, vol. 184, Article ID 115478, 2021.

[31] P. Agarwal, S. I. Hassan, S. K. Mustafa, and J. Ahmad, "An effective Diagnostic model for personalized healthcare using deep learning techniques," in *Applications of Deep Learning and Big IoT on Personalized Healthcare Services*, pp. 70–88, IGI Global, PA, USA, 2020.

[32] M. Mahrishi, S. Morwal, A. W. Muzaffar, S. Bhatia, P. Dadheech, and M. K. I. Rahmani, "Video index point detection and extraction framework using custom YoloV4 Darknet object detection model," *IEEE Access*, vol. 9, Article ID 143391, 2021.

[33] S. Bhatia, "A comparative study of opinion summarization techniques." *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 110–117, 2020.

Hindawi

Research Article

Research on Automatic Error Correction Method in English Writing Based on Deep Neural Network

Lanzhi Cheng,[1] Peiyun Ben [iD],[2] and Yuchen Qiao[3]

[1]*Zhengzhou Railway Vocational & Technical College, Zheng Zhou, Henan 450000, China*
[2]*School of Foreign Languages, Chuzhou University, Chuzhou 239000, Anhui, China*
[3]*College of Automation Science and Technology, South China University of Technology, Guangzhou 510640, Guangdong, China*

Correspondence should be addressed to Peiyun Ben; ben496668510@chzu.edu.cn

As one of the most widely used languages in the world, English plays a vital role in the communication between China and the world. However, grammar learning in English is a difficult and long process for English learners. Especially in English writing, English learners will inevitably make various grammatical writing errors. Therefore, it is extremely important to develop a model for correcting various writing errors in English writing. This can not only be used for automatic inspection and proofreading of English texts but also enable students to achieve the purpose of autonomous practice. This paper constructs an English writing error correction model and applies it to the actual system to realize automatic checking and correction of writing errors in English composition. This paper uses the deep learning model of Seq2Seq_Attention model and transformer model to eliminate deep-level errors. Statistical learning is combined with deep learning and adopted a model integration method. The output of each model is sent to the n-gram language model for scoring, and the highest score is selected as output.

1. Introduction

As one of the most widely used languages in the world, English plays a vital role in communication with the world. Now, China has become the country with the most English learners in the world. English test is an important item in the test of students' ability, and English writing ability is often the focus of the test of English proficiency. For students, due to the impact of mother tongue transfer, they often make some grammatical errors, which affect their expression. As we all know, in order to improve the level of English writing, students must do a lot of writing exercises. However, based on the teacher-student ratio, it is difficult for an English teacher to correct a large number of English compositions of each student. In this way, students cannot get timely feedback on their writing, thus failing to achieve the purpose of autonomous learning. Therefore, it is particularly necessary to develop an automatic error correction model for

English writing. With the advancement of science and technology and the development of natural language processing technology, the use of computers for automatic error correction of English writing becomes more feasible at this time [1–5].

English writing inspection includes the structure and content analysis of the upper level, as well as the spelling and sentence grammar error detection and correction. At present, there are relatively good commercial tools for word spelling error detection and correction. However, the correction of grammatical errors in students' English writing is a boring and headache problem, and there is no mature solution at present. If the computer can quickly identify writing errors in sentences and give reasonable suggestions for corrections in time, learners will have a better experience in English learning. However, some of the current commercial or open writing error detection and correction tools have relatively ordinary effects. For example, the

grammatical error checking function in Microsoft Word cannot detect common grammatical errors such as subject-verb agreement and misuse of prepositions. And the recognition rate of other errors is relatively low. The error detection rate in the automatic checking function of English writing provided by Juku.com in China is also relatively low. It is not very practical in English learning. Therefore, the breakthrough of this problem will greatly promote the application of computer-assisted English learning [6–10].

Common types of English writing grammatical errors include article errors, preposition errors, verb morphological errors, noun singular and plural errors, and subject-predicate coincidence errors. Automatic error correction in English writing is considered an extremely difficult thing, mainly for the following three reasons. (1) Many grammatical errors are context-related. They are composed of correct words. Without context, it is impossible to distinguish which word is wrong in the context. (2) The frequency of each error is relatively low, but the error modes are diverse. (3) There may be many errors in single-day sentences, which increase the difficulty of machine grammatical error correction. Early machine learning methods cannot cover complex and diverse language models, and their accuracy is low, making it difficult to achieve satisfactory results. The automatic error correction task of English writing has also ushered in new development opportunities, which reduces the influence of corpus on the task of English writing error correction to a certain extent [11–15].

Based on the deep neural network, this paper proposes a new automatic error correction model for English writing. The main contributions are as follows: first, the Seq2Seq_Attention baseline model is determined. Then, a Seq2Seq_Attention model based on the subword BPE algorithm is proposed. Next, today's hot Transformer network is used to build an automatic error correction model for English writing and introduce curriculum learning strategies and masked sequence-to-sequence strategies to improve model performance. Besides, the performance of the model is improved based on data processing and data amplification. Finally, the method of model integration is introduced to further improve the performance of the model.

2. Related Work

Rule-based methods were widely used in grammar checkers. Errors that were easier to find could be checked with simple rules, such as repeated punctuation. If an English sentence was more complicated, it needed to be checked with more complicated rules. The earlier versions of the rule-based grammar checker included EasyEnglish [16] and Park's grammar checker [17]. The most widely used tools were Microsoft Word, WordPerfect [18], and Grammarian Pro X [19], all of which could be used in multiple languages. They all used certain grammatical rules to check sentences or phrases. Therefore, the main disadvantage was that it is restricted by different languages, and a rule could not be reused by multiple languages. EasyEnglish was an English writing checker developed by IBM. It used the syntax tree represented by

the network diagram to check and find errors. It used patterns to formalize writing errors and then matched them with the constructed syntactic tree. This was a writing checker developed based on a rule-based approach and English groove grammar. However, when a sentence could not be parsed correctly and a complete grammar tree could not be constructed, the writing checker would not work well. In addition, whether it was a commercial or free tool, the grammar checker currently had no publicly released version. Park's grammar checker was a writing checker specially developed for English learners who used English as a second language to check typical errors in composition. It was a writing checker implemented by combining the rule method and the Prolog method. The writing checker could manually add rules, and new rules could replace invalid old rules. These rules could give users feedback information but would not return any information for the correct sentence.

Another method that was widely used in grammar checking was the method of statistical data analysis. Due to the rapid development of statistical machine learning methods, corpus linguistics, which used corpus as the research foundation and object, had risen rapidly. This method had been more and more recognized by researchers and had been widely used [20]. Literature [21, 22] proposed a grammatical error correction technology based on the noise channel model, which used the context of the entire sentence to correct the sentence. The basic theory of the noise channel model mainly included two parts, namely the basic language model and the noise model. The former was a probabilistic model, which generated a sentence without errors according to a given probability. The literature [23] used a variety of vocabulary and part-of-speech features, including verb phrases and nouns of adjacent prepositions, part-of-speech tags, word lemmas, to evaluate a large amount of data including newspaper news articles and English articles of intermediate and advanced English students. The error correction range covered 34 commonly used prepositions, and the final accuracy of the maximum entropy algorithm was 69%. Literature [24] used complex grammatical features, WordNet category features, and various types of part-of-speech features. At the same time, some grammatical relations extracted from the syntactic tree were also used. These grammatical relationships could be used as strong preposition checking features. Then, the data subset of the British National Corpus was used as the test set, and the correct rate was 75.6%, which was a relatively good result. Literature [25] proposed a method based on memory learning to select articles. The features he used were extracted from the Pennsylvania sentence database, for example, the part of speech of the head word, the head word of the noun phrase, and other qualifiers in the noun phrase. There were also some features extracted from the translation system from Japanese to English, such as the semantic category of the head word of noun phrases and the tendency of countability. The highest check accuracy rate obtained by the model was 83.6%. Literature [26] used the applied log-linear model to automatically recover the missing articles. It described a competitive classification model to generate and

describe the articles, thereby recovering the correct sentence. Literature [27] described a method that uses a maximum entropy classifier to select articles for noun phrases, and some features of the context are used in this process. When the training set size of the classifier reached 6 million noun phrases, the accuracy of the classification method reached 87.99%.

3. Method

This chapter first determines the baseline model of Seq2-Seq_Attention (S2SA). Then, based on data preprocessing and data amplification, the performance of the baseline model is improved. Then, a Seq2Seq_Attention model based on the BPE subword level is proposed called S2SA-BPE. Then, an English error correction model is built based on Google's Transformer model and introduced curriculum learning strategies and masked sequence-to-sequence training strategies to further improve the correction results. Finally, the N-gram language model and the deep learning model are integrated.

3.1. Error Correction Based on S2SA. This paper chooses to build an error correction baseline model based on the S2SA model for the following reasons. (1) S2SA model is a more classic model in neural network translation technology, and its position in neural network translation is equivalent to that of word2vec in the text representation. This model introduces the Attention mechanism, which breaks the restriction that the decoder can only use the encoder to finally fix the vector result. This allows the decoder to focus on the input text that is important for predicting the next target word. In addition, you can also observe the changes in the attention weight matrix to know the source input text corresponding to the target word. This helps to deepen the understanding of the model. (2) The idea of the S2SA model is relatively simple, easy to understand, and its code structure is relatively simple, which will speed up the deployment of the model. Even if the baseline model is not the final model, it can be iterated quickly, thereby reducing unnecessary time costs. (3) The baseline model is easier to deploy, generally consists of relatively few trainable parameters, and can quickly match the data without much processing. The most important thing is to facilitate research; that is, most of the errors encountered can be easier to locate the errors in the data or the defects of the model. (4) The baseline model is conducive to understanding the data, and finding out the errors in the process of constructing the baseline model is very constructive for discovering deviations and specific errors in the data. (5) The baseline model is conducive to understanding the task and helps to understand which part of the project is more difficult and which part is simpler. According to this idea, it is helpful to locate which aspect of the model should be improved, so as to better solve the difficult part. The structure of S2SA is illustrated in Figure 1.
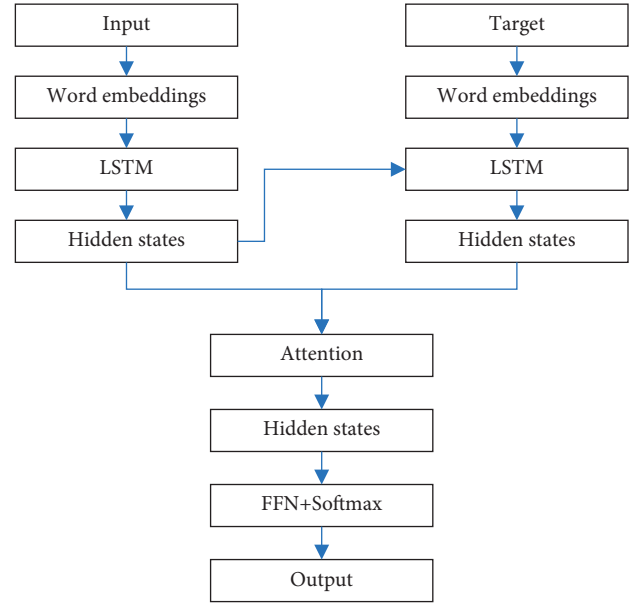
FIGURE 1: The structure of S2SA.

Encoder input is the word embedding, and the output is the hidden layer state.

$$h_t = E_{LSTM}(x_t, h_{t-1}), \tag{1}$$

where E_{LSTM} is LSTM encoder.

Decoder input is the word embedding, and the output is the hidden layer state.

$$s_t = D_{LSTM}(y_t, s_{t-1}), \tag{2}$$

where D_{LSTM} is LSTM decoder.

The context vector c_i is a weighted average of the hidden layer state.

$$c_i = \sum w_{ij} h_j. \tag{3}$$

The weight for the hidden layer state is calculated as follows:

$$w_{ij} = \frac{\exp(e_{ij})}{\sum \exp(e_{ik})}. \tag{4}$$

The weight with the hidden state is calculated as follows:

$$e_{ij} = \text{score}(s_i, h_j). \tag{5}$$

The context vector, as well as the hidden state, is concatenated as follows:

$$p_t = \tanh(W_c[c_t; s_t]). \tag{6}$$

The final output probability is calculated as follows:

$$p = \text{softmax}(W_s p_t). \tag{7}$$

3.2. Error Correction Based on S2SA-BPE. The S2SA model based on word level has the following disadvantages: (1) There are usually many words in the vocabulary that share a

vocabulary unit but have different grammatical forms. (2) There are unknown words and rare words. Unknown words refer to words that are not in vocabulary and are marked as OOV words. Rare words refer to some words in the vocabulary that appear too few times in the training corpus so that they cannot be fully trained to obtain good word vectors. (3) There is no perfect word segmentation algorithm in any language. An excellent word segmentation algorithm should be able to divide any sentence into a sequence of lexical units and grammatical forms. Therefore, this paper uses a Seq2Seq_Attention model (S2SA-BPE) based on the BPE subword level, which effectively alleviates the problem of unknown and rare word translation and improves the performance of the model. The S2SA-BPE algorithm implements a text representation unit that is between characters and words and is also different from character n-grams, achieving a more balanced state in terms of vocabulary capacity.

Byte pair encoding is a data compression method that uses an unfamiliar byte to replace the byte that frequently appears in the sentence. We use this algorithm to split words and merge characters or character sequences. The steps of the BPE learning algorithm are as follows: (1) the symbol vocabulary is initialized and added all the characters to the symbol vocabulary. Special symbols are added to the end of words. (2) All the symbols are counted, find the most frequent character pair, and replace it with a new character. (3) Each time it is merged, and a new character is generated, which means n-gram character. (4) Common n-grams characters will be merged into one symbol at the end. (5) The final symbol vocabulary size is the sum of the initial size and the number of merge operations. The number of operations is the only hyperparameter of the algorithm. The structure of S2SA-BPE is illustrated in Figure 2.

Existing word segmentation algorithms generally target normal texts and apply standard word segmentation algorithms to the task of correcting English texts. Segmentation of texts with errors will lead to mis-segmentation. Moreover, the word segmentation algorithm itself has the problem of ambiguity segmentation; that is, the word segmentation process will introduce additional error information with a high probability. In addition, the existing translation models based on word level generally limit the size of the vocabulary to alleviate the problem of excessive calculation of the softmax function. The limited vocabulary size will cause rare words to become unknown words, which will damage the performance of the model. Therefore, the S2SA-BPE method can alleviate the above problems to a certain extent.

3.3. Error Correction Based on Transformer.

3.3. Error Correction Based on Transformer. This paper builds an English writing error correction model based on the Transformer network because of the following points. (1) In terms of parallel computing capabilities, the current input of the RNN network depends on the input at the previous time, which makes it impossible to parallelize. The Transformer introduces the Attention mechanism to reduce the distance between any two characters in the text to a constant, which helps to alleviate the problem of RNN's long-distance
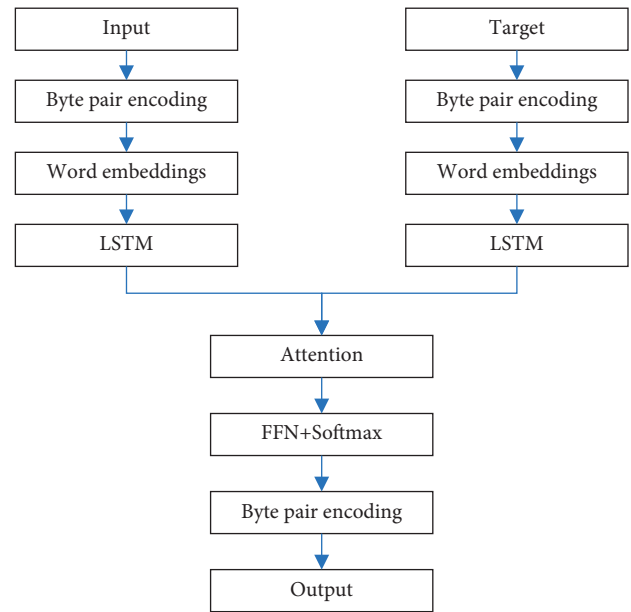
FIGURE 2: Structure of S2SA-BPE.

dependence. And due to matrix operations, its parallel computing power is also better than RNN. (2) In terms of computational efficiency, self-attention is more computationally efficient than RNN and CNN. (3) In terms of semantic feature extraction capabilities, Transformer surpasses RNN and CNN. (4) In terms of long-distance feature capture capability, the CNN feature extractor is significantly weaker than the Transformer. (5) In terms of comprehensive feature extraction capabilities represented by machine translation tasks, Transformer has a higher performance than RNN.

Based on the above five factors, this paper chooses to build a Transformer-based error correction model for English writing to further improve the effect of error correction. A big advantage of Transformer is the global receptive field; that is, RNN/CNN can only see part of the context at a time. In Transformer, each node can directly interact with other nodes. But, there are pros and cons, because Transformer does not introduce a strong prior. Therefore, a large amount of data is needed to learn a certain statistical relationship of the data from scratch. Practice has proved that the effect of Transformer on small data sets is not as good as RNN/CNN. But when there is a large amount of training data, Transformer will have a higher upper limit.

3.4. Curriculum Learning Strategy. Curriculum learning is similar to the human learning mechanism, that is, learn simple skills first and then learn difficult skills. If the training data are input in a specific order, in other words, the model first learns from simple data and then learns difficult data after the model has a certain ability, which is in line with human intuition. At the same time, from the perspective of machine learning, this method can also avoid falling into a bad local optimal solution prematurely. This can increase the generation speed and speed up the convergence speed and find a better local minimum in the nonconvex training data.

The curriculum learning method is sensitive to hyperparameters, and this paper uses a curriculum learning method with only one adjustable hyperparameter, called competence-based curriculum learning [28].

There are two crucial concepts in the learning strategy method of this course: difficulty and competence. Difficulty represents the difficulty value of a training sample. Its value is determined by the sentence length and the relative word frequency of the word. The calculation is as follows:

$$\text{diff}(s_i) = -\sum_{k=1}^{N} \log(f_k), \tag{8}$$

where s_i is the ith sample, N is the length of the sample sentence, f_k is the relative word frequency of the kth word in the sample.

Competence is a value between 0 and 1, which represents the progress of model training and is defined as a function of the model state. Specifically, this method defines the model's ability $c(t)$ at time t as the proportion of training data allowed to be used at time t. The training samples are sorted according to their difficulty, and the model only allows their top $c(t)$ part to be used at time t. Linear function and root function are two calculation methods.

$$\text{comp}_L(t) = \min\left(1, \frac{(1-c_0)t}{T} + c_0\right),$$

$$\text{comp}_R(t) = \min\left(1, \left(\frac{(1-c_0^m)t}{T} + c_0^m\right)^{1/m}\right), \tag{9}$$

where c_0 is the initial value, T is the time step threshold. When the threshold is exceeded, the model is considered to be fully capable, and t is the time step.

In this paper, the curriculum learning strategy is applied to the English writing error correction task to achieve the purpose of improving performance. The model is illustrated in Figure 3.

The competence of this paper is in addition to the linear form and root form mentioned above. A method for selecting training data based on loss is also proposed. The most intuitive reflection of the strength of the model is the loss of the model. The calculation formula is as follows:

$$\text{comp}_{\text{Loss}}(t) = \min\left(1, \left(\frac{1-\text{loss}}{T+c_0}\right)^{1/2}\right). \tag{10}$$

3.5. Masked Sequence-to-Sequence Strategy.
Most of the existing pretraining models are based on natural language understanding tasks and have achieved excellent results, which have attracted more and more attention. However, in sequence-to-sequence natural language generation tasks, such as machine translation, summary generation, automatic question and answer, few pretrained models are targeted for such tasks. Literature [29] proposed a pretraining method for natural language generation tasks: masked sequence-to-sequence pretraining (MASS). This
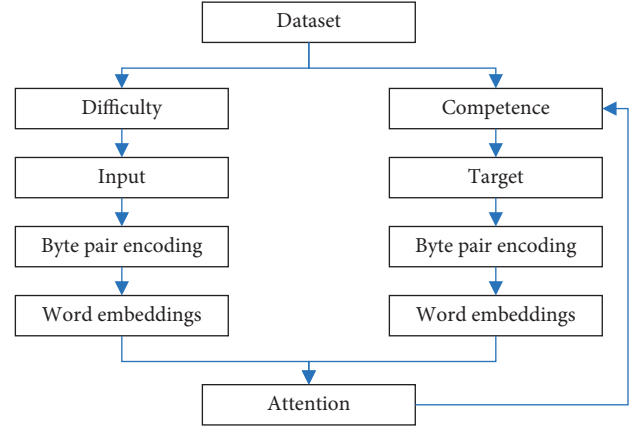
Figure 3: Based on the curriculum learning model frame diagram.

method uses the encoder-decoder framework to reconstruct a sentence segment: its encoder randomly masks multiple consecutive features of the input sentence. Then, the decoder tries to predict the features that are concealed, and its model architecture is shown in Figure 4.

MASS pretrains the encoder and decoder jointly in two steps. (1) By predicting the feature of the sentence that is masked in the encoder, MASS forces the encoder to understand the meaning of the feature of the unmasked sentence. (2) By masking the decoder input that is not masked at the source, MASS forces the decoder to rely more on the representation of the source. So as to better promote the union between encoders.

The MASS method has the following advantages: (1) The decoder side masks all words to facilitate the decoder to extract more information to improve the prediction results, thereby facilitating joint training. (2) In order to provide more information to the decoder, the encoder is forced to extract the information of the unmasked words, thereby improving the ability of the encoder to extract information. (3) Decoder is used to predict the continuous words that are obscured, which can promote the language modeling ability of the decoder. In order to improve the effect of the error correction model, this paper introduces the MASS pretraining method to the error correction task for the first time. This paper generates masking data based on char and word. It should be noted that the features are masked sequentially, not randomly.

3.6. Data Preprocessing and Data Amplification.
The reasons for data preprocessing and data amplification in this paper are as follows: (1) The higher the quality of the training data, the better the model performance. The quality of the data determines the upper limit of task performance, and necessary data preprocessing techniques can be used to approach the upper limit of task performance. (2) The more the training data, the better the model performance. Data are the core driving force of neural network models. Massive training data are one of the important reasons for the success of neural network models.
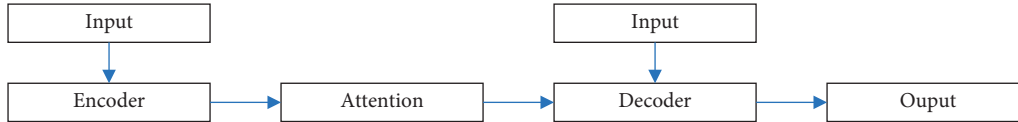
FIGURE 4: MASS method architecture diagram.

Processing the data is essential while developing a deep learning model in order to lessen the data's complication. A suitable model is established, which will help the model fit the data better and increase the pace at which it converges and improves the model's effectiveness.

Therefore, based on the above viewpoint, this paper uses edit distance to process the corpus. Edit distance [30] is also called Levenshtein distance, which is a tool to measure the similarity of two texts. Its specific meaning refers to the minimum number of editing operations required to convert one text to another. The editing here generally includes three types: inserting characters, deleting characters, and replacing characters.

In the field of natural language processing, data amplification is generally achieved in three ways. (1) Direct method searches for data related to the task and directly performs data amplification. (2) The indirect method, through the pretraining model, directly fine-tunes the trained model based on the data of its own task. This method generally requires higher hardware. (3) The method of data modification, that is, the four simple operations of synonym substitution, random insertion, random exchange, and random deletion, is used to achieve data amplification. But in the field of natural language processing, this method is rarely used, because simple addition, deletion, and modification operations can easily damage the performance of the model. In the data amplification experiment in this chapter, the direct amplification method is used to combine different corpora.

3.7. Model Integration.

Model integration is to build a series of models and uses a certain strategy to combine the built models. Then, a model with higher accuracy, better stability, and generalization effect can be obtained. At present, model integration has become a weapon in various competitions or tasks. There are many existing integration methods, but the most classic methods are Bagging and Stacking. Many methods have also been developed from this. The main idea of Bagging is to randomly select a part of the sample from the overall sample with a replacement for training and obtain multiple models by repeating the operation multiple times. Then, the average is voted or taken as the resulting output. Stacking is a layered model integration method. The input of the first layer is the original training set, and the second layer uses the output of the first layer model as the training set for training.

The model ensemble method of this paper is to learn from the Bagging method. N-gram, S2SA, S2SA-BPE, Transformer, Curriculum learning strategy-based Transformer (CL-Transformer), char-based masked sequence-to-sequence strategy Transformer (MC-Transformer), and

word-based masked sequence-to-sequence strategy Transformer (MW-Transformer) is combined for integration. The output results are scored using the N-gram language model. The highest score is used as the final output, and its multimodel integration framework is shown in Figure 5.

4. Experiment and Discussion

4.1. Experimental Environment.

The models used in this chapter are all end-to-end learning networks, and a suitable experimental environment is built according to the characteristics of the end-to-end learning network. This paper is based on the Python language and uses the PyTorch framework to code the model. The experimental environment is shown in Table 1.

4.2. Dataset and Metric.

The training forecast in this article is mainly divided into two parts. The first part comes from the International Corpus Network of Asian Learners of English (ICNALE), and the second part comes from the Brown Corpus. The testing is expected to come mainly from the learner corpus.

The F value evaluation is commonly used internationally to evaluate the effect of grammar checking. F value evaluation is also a commonly used evaluation standard in the field of natural language processing, and its calculation formula is

$$F_{\beta} = \frac{\left(1 + \beta^2\right)PR}{\beta^2 P + R},\tag{11}$$

where β is the parameter, P is precision, and R is recall rate. The calculation formula of P and R is as follows:

$$P = \frac{A}{A + B},$$
$$R = \frac{A}{A + C},\tag{12}$$

where A represents the number of sentences that actually contain grammatical errors among the sentences marked by the model as containing grammatical errors. B represents the number of sentences that do not contain grammatical errors among the sentences marked by the model as containing grammatical errors. C represents the number of sentences with grammatical errors that are not marked by the model.

This article uses $F_{0.5}$ instead of F_1, because in practical applications, we focus more on accuracy. In $F_{0.5}$, it is emphasized that the accuracy rate is twice the recall rate, and in F_1, the accuracy rate and the recall rate are equally important.
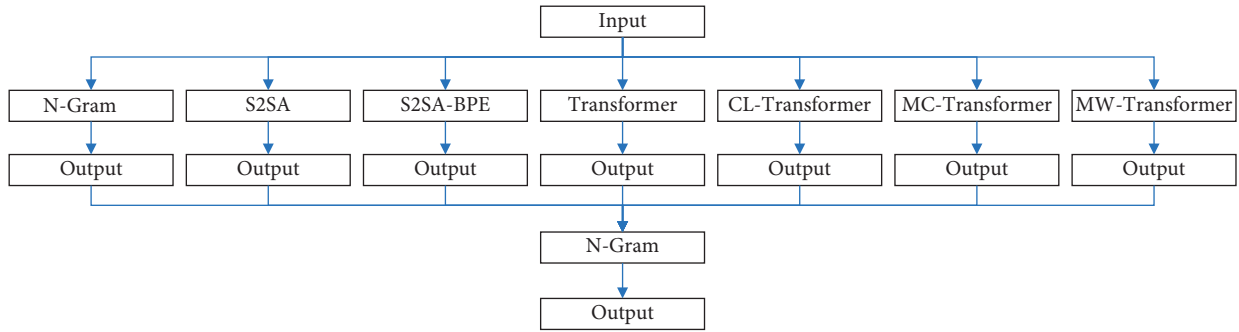
FIGURE 5: Multimodel integration framework diagram.

TABLE1: The experimental environment.

Item	Type
CPU	Intel core i7-8700K
GPU	NVIDIA GeForce RTX 3090ti
Operating system	Ubuntu 20.04
Deep learning framework	Pytorch 1.8

4.3. Evaluation on S2SA. S2SA is the baseline model of this paper. It uses word vectors to represent the input text. Both encoder and decoder use a two-layer two-way LSTM network. In order to obtain better text representation, two kinds of word vectors are used, namely Word2vec word vector and GloVe word vector. The result is illustrated in Figure 6.

The data results show that the performance of the baseline model based on the Word2vec word vector is slightly lower than that of the baseline model based on the GloVe word vector. The latter achieves 1.1%, 1.8%, and 1.6% performance improvements on three performance indicators. In subsequent experiments, GloVe word vectors are used.

4.4. Evaluation on S2SA-BPE. To solve the problem of unknown words and rare words that often appear in the word level model, BPE subword technology is introduced. The word segmentation tool used is a trained BPE subword model, and the remaining model parameters are consistent with the baseline model settings. In order to verify the effectiveness of this model, S2SA-BPE and S2SA are compared, and the experimental results are shown in Figure 7.

The data results show that under the same data and the same parameters, the performance of the model based on S2SA-BPE surpasses the performance of the model based on word level. The improvements in the three performance indicators are 1.2%, 1.8%, and 1.7%. The subword strategy can improve the performance.

4.5. Evaluation on Curriculum Learning Strategy. The curriculum learning strategy imitates the human learning mechanism, that is, learn simple knowledge first, and then learn difficult knowledge. This learning method can reduce the probability of falling into a local optimal solution and speed up the convergence speed of the model. The input of the model is that the difficulty value of the input text is less
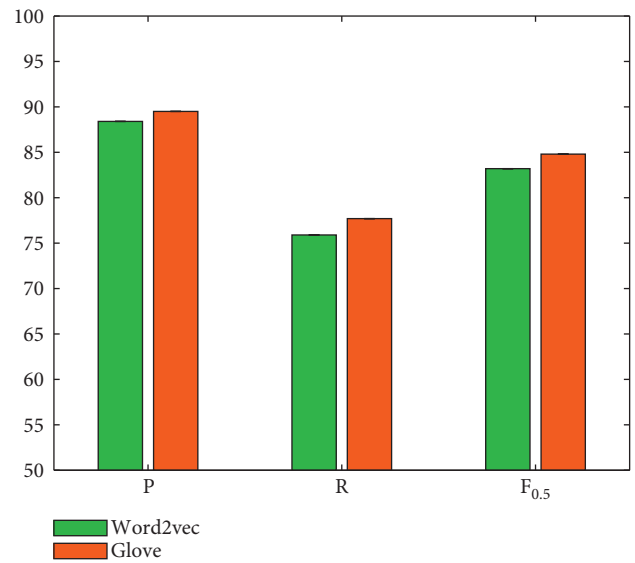
FIGURE 6: Evaluation on S2SA for different word vectors.
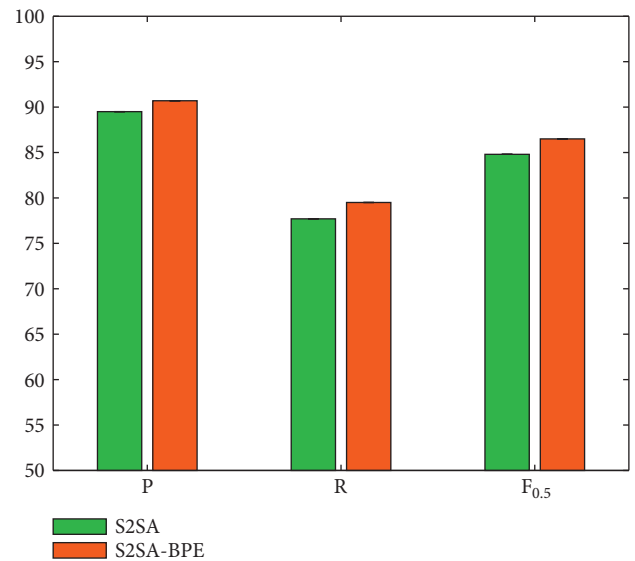
FIGURE 7: Evaluation on S2SA-BPE.

than the current ability value of the model, and three experiments are set up according to the different calculation methods of competence. They are the linear method, root

method, and loss method, and the rest are consistent with the Transformer model parameters. The result is illustrated in Figure 8.

It can be seen that curriculum learning strategies can effectively improve the performance of error correction in English writing. It should be noted that compared with the traditional Transformer method, no matter which strategy can obtain performance improvement. But compared with the linear and root methods, the loss method designed in this paper can achieve the best performance improvement. This proves the effectiveness of this design in this paper.

4.6. Evaluation on Masked Sequence-to-Sequence model.
The pretraining method is used to indirectly increase the size of the data set. The encoder and decoder are jointly trained by masking the sequence-to-sequence model, breaking the limitation that the pretraining model can only train a certain part. The model first uses the preprocessed corpus for pretraining, and the pretraining part can be divided into two parts. One is a model based on word level, and the other is a model based on word level. In order to compare these two different methods, this paper conducts a comparative experiment, and the results are shown in Figure 9.

It can be seen that the MASS strategy can effectively improve the performance of error correction in English writing. It should be noted that compared with the traditional Transformer method, no matter which strategy can obtain performance improvement. But compared with the word method, the char method designed in this paper can achieve the best performance improvement, which can achieve 0.7%, 1.1%, and 1.2% gains on precision, recall, and $F_{0.5}$.

4.7. Evaluation on Data Preprocessing and Data Amplification.
In this paper, data preprocessing and data amplification are carried out on the data used in the experiment. In order to verify the effectiveness of this data processing method, this paper conducts a comparative experiment. The performance of each model with data processing and without data processing is compared, and the experimental results are shown in Figure 10. CL-T is CL-Transformer. MC-T is MC-Transformer. MW-T is MW-Transformer.

It can be seen that after using data preprocessing and data amplification methods on each individual model, the error correction performance of the model will be improved accordingly. This experiment proves the effectiveness of the data processing method in this paper.

4.8. Evaluation on Multimodel Aggregation.
The writing error correction model designed in this article is an aggregation model, which is aggregated from S2SA, S2SA-BPE, Transformer, CL-Transformer, MC-Transformer, and MW-Transformer. In order to verify the effectiveness of this aggregation model, this paper compares the separate model with the aggregated model. The results are shown in Table 2.
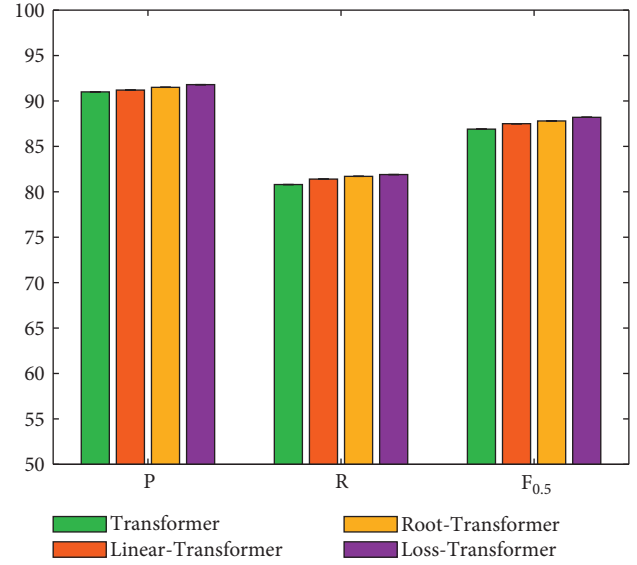
FIGURE 8: Evaluation on curriculum learning strategy.
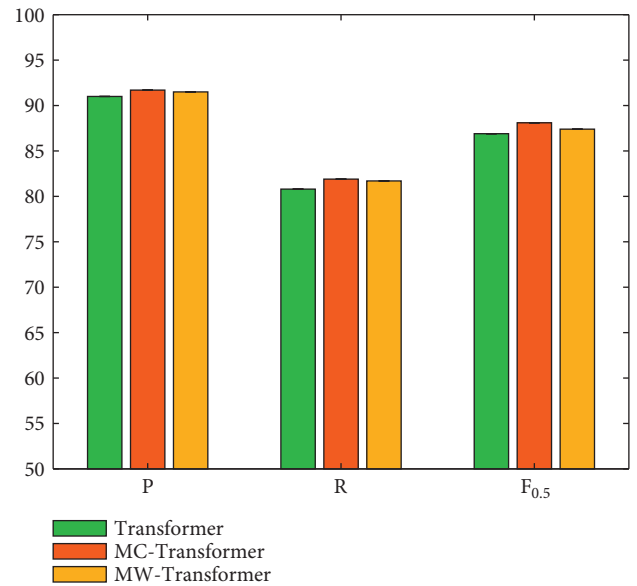
FIGURE 9: Evaluation on masked sequence-to-sequence strategy.

The data show that the aggregation model can achieve the best performance compared to each individual model. This proves the reliability and correctness of the error correction algorithm for English writing based on the aggregation model proposed in this paper.

4.9. Comparison with Other Methods.
In order to further verify the effectiveness of the method in this paper, the aggregation model is compared with other English writing error correction methods. The compared methods in this work include the naive Bayesian model (NBM), decision tree model (DTM), maximum entropy model (MEM), and KNN. The result is illustrated in Table 3.
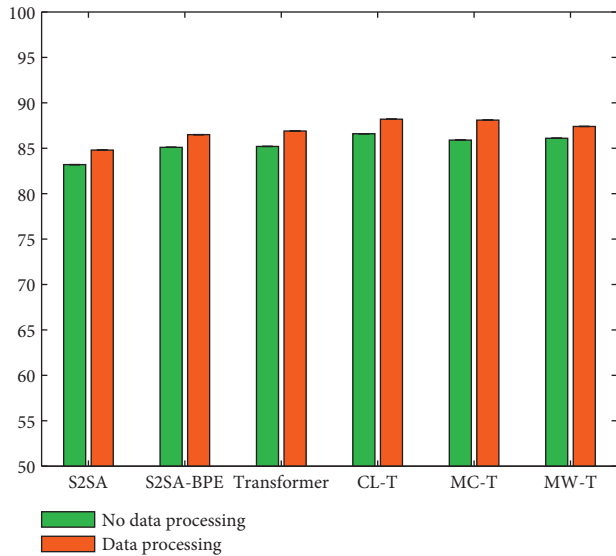
Figure 10: Evaluation on data preprocessing and data amplification.

Table 2: Evaluation on multimodel aggregation.

Model	P	R	$F_{0.5}$
S2SA	89.5	77.7	84.8
S2SA-BPE	90.7	79.5	86.5
Transformer	91.0	80.8	86.9
CL-transformer	91.8	81.9	88.2
MC-transformer	91.7	81.9	88.1
MW-transformer	91.5	81.7	87.4
Ours	92.5	84.3	89.5

Table 3: Comparison with other methods.

Method	P	R	$F_{0.5}$
NBM	79.8	68.1	73.5
DTM	84.3	72.5	77.8
MEM	87.5	81.4	84.3
KNN	89.3	82.1	85.8
Ours	92.5	84.3	89.5

Compared with other error correction methods, the aggregation model method designed in this paper can obtain the best performance. Compared with the best KNN method listed, the three performance indicators have been improved by 3.2%, 2.2%, and 3.7%, respectively.

5. Conclusion

With the development of artificial intelligence and computer science and technology, natural language processing technology has developed rapidly, providing a theoretical and technical basis for the intelligent assistance of English. A large number of teaching and tutoring software has emerged in the following. As an important part of English learning, writing has become a difficult challenge for learners in the learning process. For students, the strength of the ability to use English writing directly affects the level of English proficiency. At the same time, this will also affect English reading and speaking skills. Based on the deep neural network, this article is dedicated to developing an intelligent model for correcting various errors in English writing. This can not only be used for automatic inspection and proofreading of English writing but also enable students to achieve the purpose of autonomous practice. First of all, this paper determines the S2SA baseline model. Then, an S2SA-BPE model based on the subword BPE algorithm is proposed. Afterward, we used the now hot Transformer network to build an error correction model for English writing. And the curriculum learning strategy and the masking sequence-to-sequence strategy are introduced to improve the performance of the model. Then, the model performance based on data processing and data amplification is improved. Finally, the method of model integration is introduced to efficiently aggregate the various submodels designed. This can further improve model performance.

Data Availability

The datasets used are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] C. Y. Jie, C. H. Chen, and M. Changjeng, "Integrating video-capture virtual reality technology into a physically interactive learning environment for English learning," *Computers & Education*, vol. 55, no. 3, pp. 1346–1356, 2010.

[2] T.-Y. Kim, "Korean elementary school students' English learning demotivation: a comparative survey study," *Asia Pacific Education Review*, vol. 12, no. 1, pp. 1–11, 2011.

[3] J. M. Carreira, K. Ozaki, and T. Maeda, "Motivational model of English learning among elementary school students in Japan," *System*, vol. 41, no. 3, pp. 706–719, 2013.

[4] C. C. Chang, C. F. Yan, and J. S. Tseng, "Perceived convenience in an extended technology acceptance model: mobile technology and English learning for college students," *Australasian Journal of Educational Technology*, vol. 28, no. 5, pp. 809–826, 2012.

[5] A. I. Fageeh, "EFL learners' use of blogging for developing writing skills and enhancing attitudes towards English learning: an exploratory study," *Journal of Language and Literature*, vol. 2, no. 1, pp. 31–48, 2011.

[6] S.-W. Yeh, J.-J. Lo, and H.-M. Chu, "Application of online annotations to develop a web-based Error Correction Practice System for English writing instruction," *System*, vol. 47, pp. 39–52, 2014.

[7] M. Chodorow, M. Gamon, and J. Tetreault, "The utility of article and preposition error correction systems for English

language learners: feedback and assessment," *Language Testing*, vol. 27, no. 3, pp. 419–436, 2010.

 [8] A. A. Khansir and F. Pakdel, "Place of error correction in English language teaching," *Educational Process: International Journal*, vol. 7, no. 3, pp. 189–199, 2018.

 [9] W. Lu, "Differences between contrastive analysis and error analysis on English writing for Chinese EFL learners," *Overseas English*, vol. 396, no. 08, pp. 277-278, 2019.

[10] Y. Wang, "Functions, values & inadequacies-an evaluative discussion of pigai intelligent online English writing correction system in view of second language acquisition," *Journal of Physics: Conference Series*, vol. 1237, Article ID 042002, 2019.

[11] A. Ariyanti, "The effectiveness of indirect written correction on English writing skills," *Journal of English for Academic and Specific Purposes*, vol. 1, no. 2, p. 33, 2018.

[12] J. H. Lee, M. Kim, and H. C. Kwon, "Deep learning-based context-sensitive spelling typing error correction," *IEEE Access*, no. 99, p. 1, 2020.

[13] Y. Bassil and M. Alwani, *OCR Context-Sensitive Error Correction Based on Google Web 1T 5-Gram Data Set[J]*, Arxiv: 1204.0188, 2012.

[14] R. De Felice and S. Pulman, "Automatic detection of preposition errors in learner writing," *Calico Journal*, vol. 26, no. 3, pp. 512–528, 2009.

[15] Y. Z. Chen, S. H. Wu, and P. C. Yang, "Improve the detection of improperly used Chinese characters in students' essays with error model," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 21, no. 1, pp. 103–116, 2011.

[16] W. Xiang and P. Jin, "Performance comparison of stanford parser and berkeley parser based on large corpus," *Computer Knowledge and Technology*, vol. 9, no. 8, pp. 1984–1986, 2013.

[17] M. R. Cavaleri and S. Dianati, "You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students," *Journal of Academic Language and Learning*, vol. 10, no. 1, pp. A223–A236, 2016.

[18] G. C. Allen, S. M. Rodman, and B. M. McGrew, "WordPerfect: office applications[J]," *Otolaryngology-Head and Neck Surgery*, vol. 112, no. 5, p. 146, 1995.

[19] Y.-L. Wu, *The Impact of Technology on Language Learning [M]. Future Information Technology*, Springer, Berlin, Heidelberg, 2014.

[20] L. Aversano, G. Canfora, A. De Lucia, and S. Stefanucci, "Evolving ispell: a case study of program understanding for reuse," in *Proceedings of the International Workshop on Program Comprehension*, pp. 197–206, IEEE, Paris, France, June 2002.

[21] Y. A. Park and R. Levy, "Automated whole sentence grammar correction using a noisy channel model," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 152–165, Portland Oregon, June 2011.

[22] R. West, Y. Alet Park, and R. Levy, "Bilingual random walk models for automated grammar correction of ESL," in *Proceedings of the Author-Produced Text Proceedings of the Sixth Workshop on Innovative Use of NLP for Building Educational Applications*, pp. 170–179, California, USA, June 2011.

[23] M. Chodorow, J. Tetreault, and N. R. Han, "Detection of grammatical errors involving prepositions," in *Proceedings of the ACL-SIGSEM Workshop on Prepositions*, pp. 25–30, Prague Czech Republic, June 2007.

[24] R. De Felice and S. Pulman, "Automatically acquiring models of preposition use," in *Proceedings of the ACL-SIGSEM Workshop on Prepositions*, pp. 142–150, Prague Czech Republic, June 2007.

[25] G. Minnen, F. Bond, and A. Copestake, "Memory-based learning for article generation," in *Proceedings of the Conference on Computational Language Learning and the 2nd Learning Language in Logic Workshop*, pp. 43–48, Lisbon, Portugal, September 2000.

[26] S. Bischoff, D. Pavic, and L. Kobbelt, "Automatic restoration of polygon models," *ACM Transactions on Graphics (TOG)*, vol. 24, no. 4, pp. 1332–1352, 2005.

[27] N.-R. Han, M. Chodorow, and C. Leacock, "Detecting errors in English article usage by non-native speakers," *Natural Language Engineering*, vol. 12, no. 2, pp. 115–129, 2006.

[28] E. A. Platanios, O. Stretcu, G. Neubig, B. Poczos, and T. M. Mitchell, "Competence-based curriculum learning for neural machine translation," 2019, https://arxiv.org/abs/1903.09848.

[29] K. Song, X. Tan, T. Qin, J. Lu, and T. Y. Liu, "Mass: masked sequence to sequence pre-training for language generation," 2019, https://arxiv.org/abs/1905.02450.

[30] G. D. Bergland, "A guided tour of the fast Fourier transform," *IEEE spectrum*, vol. 6, no. 7, pp. 41–52, 1969.

Research Article

New Progress in Artificial Intelligence Algorithm Research Based on Big Data Processing of IOT Systems on Intelligent Production Lines

He Shi,[1] Guohua Cao ®,[1,2] Guoqing Ma,[1,3] Jingsong Duan,[1] Jimeng Bai,[1] and Xiangyin Meng[4]

[1]*College of Mechanical and Electrical Engineering, Changchun University of Science and Technology, Changchun 130022, Jilin, China*
[2]*Chongqing Research Institute of Changchun University of Science and Technology, Chongqing 401135, China*
[3]*Wuhu HIT Robot Technology Research Institute CO., LTD, Wuhu 241000, Anhui, China*
[4]*College of Mechanical Engineering, Southwest Jiaotong University, Chengdu 611756, Sichuan, China*

Correspondence should be addressed to Guohua Cao; caogh@cust.edu.cn

Intelligent production line is the abbreviation of intelligent production line. Intelligent production line refers to a form of production organization that uses intelligent manufacturing technology to realize the production process of products. The actual manufacturing process includes different levels and links, and each step cooperates to create a high-efficiency production line. The intelligent production line includes 3 levels covering automation equipment, digital workshops, and intelligent factories and runs through 6 major links of intelligent manufacturing (intelligent management, intelligent monitoring, intelligent processing, intelligent assembly, intelligent inspection, and intelligent logistics). The emergence of the Internet of Things system has changed the way of information dissemination. The system combines radio frequency automatic identification and global positioning system technologies to achieve functions such as information exchange and processing, enabling information processing to be intelligent and improving resource utilization. Big data processing includes multiple data processing procedures, but data quality is the most important link in the entire process, and each data processing link will have an impact on the quality of big data. The big data processing process mainly includes data collection, data preprocessing, data storage, data processing and analysis, data display, data visualization, data application, and other links. This article aims to study the new progress of artificial intelligence algorithms for big data processing of IOT systems on intelligent production lines. It is hoped that through the development of intelligent production lines and big data processing technologies, ways to optimize artificial intelligence algorithms can be found. This study proposes a metadata replication method based on a separate replication strategy, which separates the replication process of the data operation log, each is independent, and shortens the data replication time. Combining the existing intelligent production line network platform in the laboratory and carrying out the research of the intelligent production line network state prediction system based on the neural network to design a network prediction system can prejudge the operation status of the intelligent production line network. The experimental results in this article show that when the Namenode mode is used to read data and when the number of clients reaches 8, the data processing basically remains unchanged. When the NCluster system reads data and when the number of clients is 6, the data is processed 1256. When the number of clients is 20, the data is processed 2100, the NCluster system will remain stable when the number of clients reaches 12, and compared with the Namenode system, it has obvious advantages.

1. Introduction

The scope of artificial intelligence applications is very wide, including medicine, diagnosis, financial trade, robot control laws, scientific discoveries, and toys. Artificial intelligence, also known as AI, has produced many aspects to solve the most difficult problems in computer science. However, with the ever-expanding social production field, people are increasingly raising production standards, which urgently requires the continuous optimization of artificial intelligence algorithms to reach higher standards. As an important part of the Internet, the Internet of Things can form a new ecosystem of the physical world and the information world, further realize the information sharing of people, machines, and things, make the information exchange between the information world and the physical world closer, and improve social production efficiency and meet people's growing needs. With the increasing popularity of Internet technology, the information data in life have increased exponentially, so the concept of big data came into being. Big data refers to the amount of data involved is so large that cannot be retrieved, managed, processed, and organized in a reasonable time through mainstream software tools to help enterprises make business decisions more positively. How to find the required information from the massive data has become a major challenge for big data processing. With the continuous optimization of big data processing technology, people have been able to select the required information by keywords, which also improves work efficiency and saves time. At present, the development time of artificial intelligence is relatively short, and the algorithm has not been perfected. How to optimize calculation with the help of big data processing and the Internet of Things technology has become a current research hotspot.

Although artificial intelligence is still under development, its future still has many uncertainties, but the rise of artificial intelligence will promote industrial changes and profoundly affect people's lifestyles. There has been a broad consensus that this technology will definitely be a technological revolution with far-reaching impact, so the exploration of artificial intelligence is very necessary. Artificial intelligence technology will have a profound impact on the country's international competitiveness, change the pattern of international industrial division of labor, and reshape the development of the world economy. And focusing on the research of artificial intelligence technology is an important way to improve international competitiveness. In daily life, artificial intelligence technologies such as deep learning, image recognition, and voice recognition have been widely used in smart terminals, smart homes, mobile payments, and other fields to provide ordinary people with wider coverage, better experience, and better convenience in life service.

In view of the characteristics of mass, this article summarizes the related work of the cloud-centric the Internet of Things and summarizes and analyzes the current solutions and research status for the massive data of the Internet of Things. In view of the current research status of the Internet of Things data processing at home and abroad and the analysis of the demand for complex event processing of the

Internet of Things, the data processing technology system architecture of the Internet of Things is proposed. Introducing an intelligent recommendation model and using a knowledge base to provide users with a scientific algorithm recommendation plan through the ability of rough set theory to classify knowledge can realize the intelligent recommendation of the algorithm.

2. Related Work

Although artificial intelligence technology is a current hot research field, it has a short development time and is not mature enough to integrate with the production field to achieve high-efficiency production. Therefore, how to optimize artificial intelligence algorithms has become a current problem. Yi-Yu systematically proposed the strategies, methods, and engineering theories of intelligent lean manufacturing of traditional Chinese medicine in response to the "six scientific problems" and "five technical challenges" of intelligent manufacturing and lean production of traditional Chinese medicine. The holistic view of traditional Chinese medicine draws on the concept of international advanced pharmaceutical technology and applies the theory and method to six traditional Chinese medicines including Sanqi Panax Notoginseng injection. Based on this theory and method, a number of intelligent production lines have been designed and built, which has greatly accelerated the digital, networked, and intelligent manufacturing of CM. In summary, this theory and application provide a technical demonstration for the technological upgrading and high-quality development of the CM industry [1]. The Internet of Things has a significant impact on the development of manufacturing technology. Therefore, based on the analysis of the challenges and opportunities faced by the manufacturing industry, Liu took the assembly process of mechanical products as the research object and analyzed the characteristics of the manufacturing system based on the Internet of Things. In order to improve the interconnection, perception, efficiency, and intelligence of the assembly system, the author proposed the concept of an intelligent assembly system for mechanical products based on the Internet of Things. Then, introduced the IIASMP framework based on advanced technologies such as information and communication technology, sensor networks, and radio frequency identification. The key technologies such as assembly resource identification, information interaction technology, multisource data perception and fusion, intelligent assembly agent, and data value-added and dynamic adaptive optimization under this framework are described [2]. Wang proposed a new hybrid method that uses an integrated data fluctuation network and multiple artificial intelligence algorithms, called the DFN-AI model. In the proposed DFN-AI model, the author uses complex network time series analysis technology as the preprocessor of the original data to extract fluctuation characteristics and reconstruct the original data, and then uses artificial intelligence tools to model the reconstructed data and predicts the future data. To verify these results, the author checked daily, weekly, and monthly price data from the Cushing Crude Oil

Trading Center in Oklahoma. The empirical results show that the proposed DFN-AI model is significantly better than its corresponding single AI model in terms of prediction direction and prediction level. This confirms the effectiveness of the proposed nonlinear model modeling hidden in crude oil prices [3]. Bao uses the Java programming language to design and develop an RPO analysis software based on artificial intelligence (AI) algorithms. The software integrates five different AI algorithms, which can be used for RPO solutions in different situations. Users can quickly select the appropriate RPO calculation algorithm according to the actual situation, and the calculation results can guide the grid company to optimize parameters such as reactive power compensation capacity and transformation ratio. The software mainly includes a project management module, data editing module, RPO calculation module, and control variable editing module. It has a BPA-Matpower data interface, compatible with BPA data format, and can complete the accurate conversion of BPA data to Matpower data. The software can directly edit the BPA data of different grid structures without reprogramming and has the characteristics of high precision and strong versatility [4]. Dou proposed an independent regional connection model in the context of sensor networks to ensure global connectivity and satisfactory data service quality. The author also studied the optimization of deterministic and random deployment of sensor coverage and regional connections in industrial IOT systems. First, a new optimal network that achieves full sensing coverage and ensures regional connectivity is proposed for deterministic deployment. In addition, the optimal model is derived and the advantages of the proposed model are analyzed. Secondly, based on the assumption that a given sensor is deployed as a Poisson point process, a theoretical analysis is proposed to determine the minimum number of sensors for random deployment to achieve a certain coverage and connectivity. Numerical results show that the model is effective for the application of sensor networks in industrial IOT systems [5]. IOT devices are usually small, low-cost, and limited in resources, which makes them vulnerable to physical, side-channel, and cloning attacks. In order to solve this problem, Aman proposed a lightweight mutual authentication protocol based on the physical unclonable function of the Internet of Things system. Two scenarios of protocols are provided, one is when the IOT device and the server want to communicate, and the other is when two IOT devices want to establish a session. The security and performance analysis of the protocols show that they can not only resist different types of attacks but are also very efficient in terms of computing, memory, energy, and communication overhead [6]. Kim proposed an enhanced Access Reservation Protocol (ARP) with a Partial Preamble Transmission (PPT) mechanism for the narrow band Internet of Things (NB-IOT) system. The proposed ARP can improve ARP performance by reducing the occurrence of preamble collisions and is compatible with traditional NB-IOT ARP. At the same time, an analysis model is provided that can capture the performance of the proposed ARP in terms of false alarms, false detections, and collision probability. In addition, the author studied the trade-off between

false detection and collision probability and optimized the proposed ARP based on the system load. The results show that the proposed ARP is better than the traditional NB-IOT ARP, especially under heavier system load [7]. Device capture attacks usually threaten the realization of signcryption on unattended devices by extracting encryption keys from the captured devices. Inspired by this problem, Yang proposed a novel and specialized obfuscable aggregatable signcryption scheme (OASC) and signcryption algorithm obfuscator. The design of the scheme considers that the computational and communication costs should be small enough (lightweight), in order to adapt to the applications in resource-constrained embedded devices. The proposed obfuscator can protect the signcryption program from key extraction attacks by converting the program into an incomprehensible obfuscation program [8]. Although these theories have explored artificial intelligence algorithms and big data processing technology to a certain extent, the two fields are less integrated and not practical.

3. New Development Methods of Artificial Intelligence Algorithms for Big Data Processing in IOT Systems on Intelligent Production Lines

3.1. Big Data Processing. Big data brings many problems to data processing. Taking traffic security data as an example, after years of informatization construction, public security, and government departments have collected massive amounts of traffic security-related data resources, the scale and complexity of which have exceeded the management and processing capabilities of traditional databases and software technologies. Since the Internet of Things technology has been promoted, big data is everywhere in daily life, and the scope of application continues to expand [9, 10]. With the emergence of the scale of big data, it is urgently required to classify and process big data. In this context, a big data system has emerged [11], and the structure of the big data system is shown in Figure 1:

The big data processing system mainly includes data integration layer, data storage layer, data processing layer, data analysis and application layer, and platform management, and other components. The data integration layer is responsible for adapting between external data sources and the internal data storage layer of the big data processing system to achieve efficient data import and export.

With the development of sensor technology, more and more data are collected, prompting the Internet of Things to realize the concept of universal connection to connect every item at any time and place [12, 13]. Using sensors of different specifications to collect data has different data formats, so there will be heterogeneous data, which requires the data to be preprocessed before the data is analyzed [14, 15]. From the original database to the mining database, the operation on the data is called data preprocessing. Preprocessing of data is a complicated process, and different processing programs need to be written for different situations [16]. The data preprocessing structure diagram is shown in Figure 2.
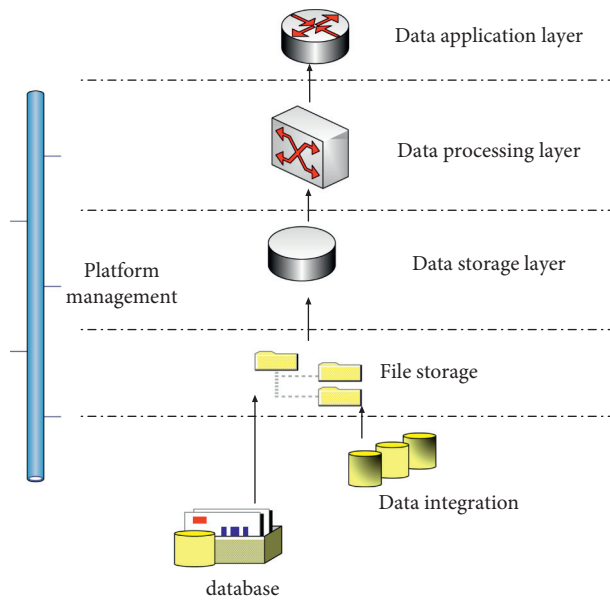
FIGURE 1: Big data system structure.
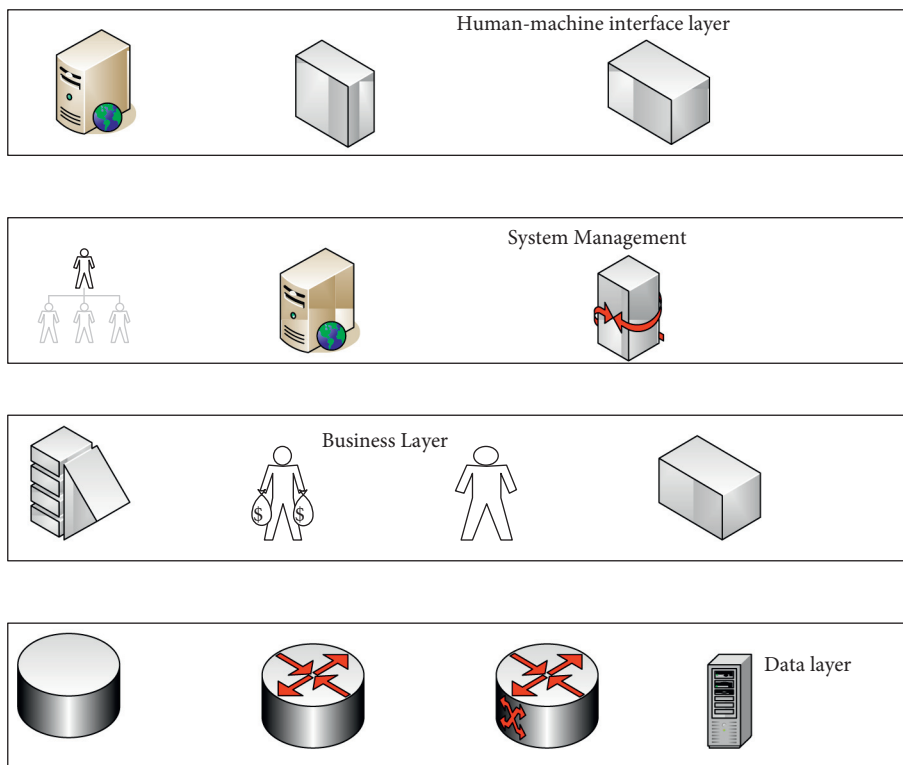
FIGURE 2: Data preprocessing structure.

3.2. Artificial Intelligence Algorithm. This section mainly discusses related theories about artificial intelligence algorithms, deepens the understanding of artificial intelligence algorithms through the discussion of related theories, and prepares for in-depth discussion [17]:

$$y(a \mid c) = \frac{y(c \mid a) y(a)}{y(c)} = \frac{y(c \mid a) y(a)}{\int y(c \mid a) y(a) da}, \quad (1)$$

where $y(c \mid a)$ represents the likelihood function, and $y(a)$ represents the probability density.

$$y(c_1 \mid a_2) = \int y(c_1 \mid c_2)dc_2. \tag{2}$$

Formula (2) represents a functional expression for predicting the current probability density function.

$$y(c_1 \mid a_2) = \frac{y(c_1 \mid a_1)y(c_1 \mid a_2)}{\int y(c_1 \mid a_1)y(c_1 \mid a_2)da_1}. \tag{3}$$

Formula (3) represents multiple integral operation.

$$W = \int_S H(a)da, \tag{4}$$

where S is a random area in the space, and W is a multiple integral.

$$W = \int_S h(a)y(a)da. \tag{5}$$

In formula (5), $y(a) \geq 0$ and $\int_S y(a)da = 1$.

$$h_k = \frac{1}{k}\sum_1^k h(a_1). \tag{6}$$

Formula (6) represents the arithmetic average value.

$$Q(1) = Q(h(a)) = \frac{1}{k}\sum_1^k h(a_1), \tag{7}$$

where $h(a)$ is an unbiased estimate of Q.

$$Q(j(a)) = \int j(a)y(a)da, \tag{8}$$

$$Q(j(a)) = \int j(a)\alpha(a)\beta(a)da. \tag{9}$$

Among them, $\beta(a)$ represents the density function, and $Q(j(a))$ represents the expected estimate.

$$Q(j(a)) = \frac{1}{l}\sum_1^l j(a)\beta(a), \tag{10}$$

where $\beta(a) = y(a)/w(a)$.

$$w(a) = w_1(a_1)w_2(a_2 \mid a_1)\cdots w_c(a_c \mid a_1 \cdots a_c), \tag{11}$$

where a represents a scalar or a vector.

$$y(a) = y_1(a_1)y_2(a_2 \mid a_1)\cdots y_c(a_c \mid a_1 \cdots a_c). \tag{12}$$

In the actual measurement process, each step is independent of each other, so formula (12) represents the density expression of the objective function.

$$\beta(a) = \frac{w(a) = w_1(a_1)w_2(a_2 \mid a_1)\cdots w_c(a_c \mid a_1 \cdots a_c)}{y(a) = y_1(a_1)y_2(a_2 \mid a_1)\cdots y_c(a_c \mid a_1 \cdots a_c)}. \tag{13}$$

Formula (13) represents the importance weight.

$$\beta_j(a_j) = \beta_{j-1}(a_{j-1})\frac{y_j(a_{j-1})}{w_j(a_{j-1})}, \tag{14}$$

$$y(a_j \mid g_j) = \sum_1^h \beta_j^l \iota(a_j - a_j^l), \tag{15}$$

where $y(a_j \mid g_j)$ stands for posterior probability.

$$\beta_j^l = \beta_{j-1}^l \frac{y(g_j \mid a_j^l)y(a_j^l \mid a_{j-1}^l)}{w(a_j^l \mid a_{j-1}^l, g_j)}. \tag{16}$$

Formula (16) represents the approximate posterior probability density of sampled particles.

$$y(a \mid g) = \sum_1^h \beta_k^h \iota(a_k - a_k^h), \tag{17}$$

where $y(a \mid g)$ stands for prior probability.

$$\beta_j^l = \frac{y(a_j^h \mid g_k)}{w(a_j^h \mid g_k)}. \tag{18}$$

Formula (18) represents the expression of particle weight function.

$$y(a \mid g) = \frac{y(g_j \mid a_{1-j}, g_{j-1})y(a_{1-j} \mid g_{j-1})}{y(g_j \mid g_{j-1})}. \tag{19}$$

Formula (19) represents the expression of the posterior probability density function.

$$y(a_j \mid a_{j-1}, g_j) = y(a_j \mid a_{j-1}). \tag{20}$$

Formula (20) represents the importance density function in the particle filter algorithm.

3.3. IOT System.

The Internet of Things is a comprehensive information processing technology based on the network and serving the purpose. It not only makes the way of communication between people leap forward but also makes it possible to communicate between people and things and between things and things. In a word, the Internet of Things technology has turned the whole world into a whole [18, 19]. Networking, IOT, interconnection, automation, perception, and intelligence are the basic characteristics of the Internet of Things. Although the function of the Internet of Things is very powerful, the foundation of the Internet of Things is still the Internet, but the Internet of Things is expanded on the basis of the Internet [20–23]. From the current research progress, the Internet of Things can be divided into perception layer, application layer, and network layer, as shown in Figure 3:

The job of the recognition layer is to recognize the target and record the corresponding information [23]. In this
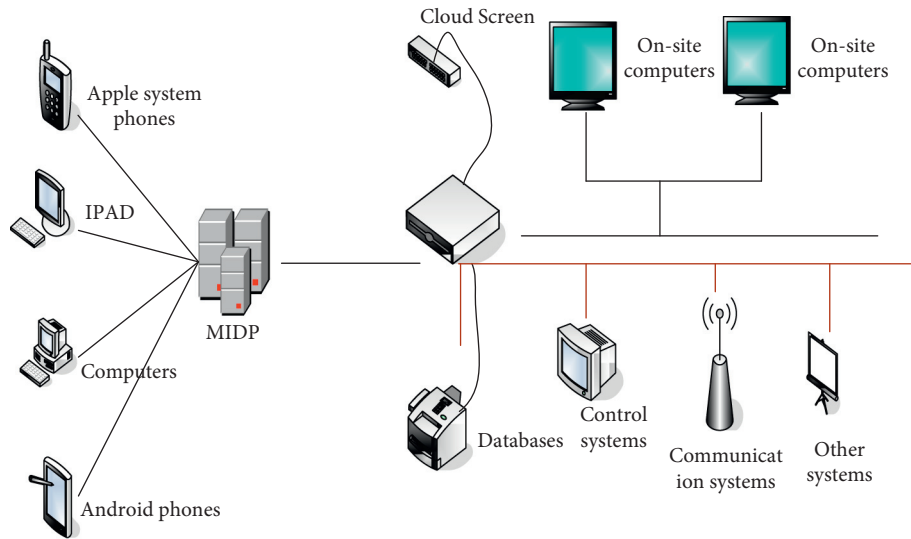
Figure 3: Internet of Things platform architecture diagram.

process, network sensor technology or recognizer equipment is involved, and only when multiple technologies cooperate with each other can accurate recognition of the target be achieved [24, 25].

The transportation layer is actually the control backbone of the Internet of Things, which is responsible for the transmission and processing of data information. It includes various corresponding communication networks and networks integrated with the Internet, which are relatively mature parts, such as the Internet, radio and television networks, and so on [26, 27].

The application layer is composed of data collection, monitoring, and energy management. The data collection and monitoring system is to monitor and control the equipment with the computer as the core in order to understand the various data in the system. The management system involves statistics and management functions. All in all, the application layer is the technology that combines the Internet of Things technology with other fields to obtain new intelligent solutions.

The Internet of Things is a product of the continuous development of information technology. Information technology was first used in the telegraph and telephone. With the continuous innovation of science and technology, the Internet of Things technology has also been continuously developed. The traditional communication structure of the Internet of Things is shown in Figure 4:

The Internet of Things technology is now regarded by workers as the core of the information industry. In the field of the Internet of Things, various elements can communicate with each other and only need to realize information sharing through the Internet of Things technology. The data in the Internet of Things have the characteristics of heterogeneity, mass, and real time. And it has the following characteristics:

(1) Transmissibility. The Internet of Things technology can realize the transmission of information by relying on relevant networks and complying with the corresponding network protocols.

(2) Comprehensive perception. The Internet of Things is a collection of perception abilities for things and people, it uses technologies such as QR codes to extract information, and it has organizational capabilities, which breaks the dependency of traditional network technologies.

(3) Automatic control. Through the use of fuzzy recognition and other corresponding control technologies, related objects can be intelligently managed and used.

(4) Intelligent processing. The intelligent processing of the Internet of Things refers to the technology that uses the corresponding data processing technology to process the native data and then transmit it to the user.

With the continuous development of science and technology, wireless communication technology has become increasingly mature, and monitoring technology has also been widely used. The early monitoring system used instrument monitoring, followed by a computer monitoring system, which consisted of controlled objects, detection parts, and execution devices. In actual production, some equipment is becoming more and more dispersed, and the independent monitoring of each equipment can no longer keep up with the requirements of monitoring, so a distributed control system appears. The transmission provides centralized monitoring that may follow the needs of social production. Traditional monitoring technology has been unable to adapt, so a distributed control system appears and wireless sensors are produced under this situation. In actual use, multiple wireless sensors are combined to form a wireless sensor network, multiple sensors are placed in the monitoring area to monitor the targets that need to be monitored in real time, and finally, the collected information is systematically analyzed. The core principle of wireless sensor network is that in a specific environment, certain nodes require higher energy, stronger communication
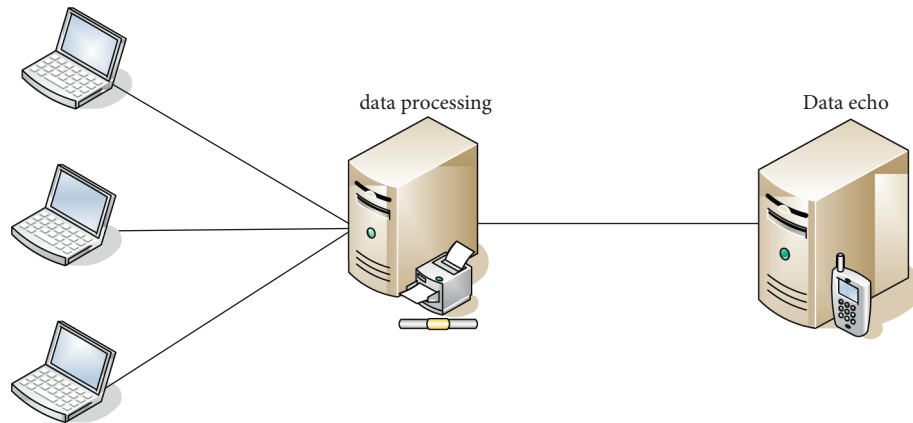
Figure 4: The traditional communication structure of the Internet of Things.

distance, and stronger processing capabilities, even including GPS positioning and other functions, and can temporarily act as cluster heads or base stations to provide the function of external network connection. Its specific structure is shown in Figure 5:

By randomly spreading or manually deploying a large number of sensor nodes in the monitoring area, real-time monitoring and collection of the sensing objects that observers are interested in, and then using the node embedded processing system to analyze and process the collected information. The data collected by the nodes in the cluster are aggregated to the cluster head node and then routed to the base station in multihop communication after the cluster head node data is merged. The base station finally transmits the data to the user terminal through wireless or wired means, realizing the user to the remote real-time monitoring of targets.

Currently in the era of information explosion, how to quickly obtain effective information is the focus. Wireless sensors can obtain a large amount of information by virtue of their distributed characteristics and are widely used in various fields of social production. Wireless network sensors were first used in the military field. Because of their strong concealment, they can play a role on the battlefield. They can monitor and evaluate the situation on the battlefield and can provide a scientific basis for military strategy decisions. In addition to outstanding performance in the military field, wireless sensors have also fully demonstrated their functions in the medical field. The tiny sensor is installed in the patient's body to monitor the patient's physical condition in real time. In case of emergencies, it can be discovered in time for rescue time. Environmental experts can use the data collected by sensor nodes to study the ecological environment to monitor and track the migration of migratory birds and wild animal life, to monitor the forest environment and fire warning, as well as problems that occur during the mining of mineral resources, such as gas concentration monitoring and water seepage conditions.

4. New Progress Experiment on Artificial Intelligence Algorithms for Big Data Processing of IOT Systems on Intelligent Production Lines

4.1. Experimental Hardware. This study discusses the operation of the Internet of Things system in artificial intelligence algorithms, especially when the Internet of Things technology is needed. Different specifications of the system hardware will also have different effects on the experiment. Table 1 is the hardware parameters of this experiment.

According to the data in Table 2, the storage of the server used in this experiment is 16 TB, the RAM is 64 GB, the network card standard is 2 Gbps, the CPU is 4 cores, and the hard disk is 2 TB.

4.2. Detector Performance Parameters. There are many different types of artificial intelligence algorithms, in order to find the best intelligent algorithm, we need to detect and analyze each algorithm. Table 3 lists the comparison data of specific artificial intelligence algorithms.

4.3. Simulation Environment Parameters. Different algorithms have different functions. In the specific experiment process, performance analysis and evaluation of different types of algorithms are required to ensure the authenticity of the data. The simulation parameters used in this article are all obtained based on the data of real nodes in the IOT experimental platform. The specific conditions are as follows:

According to the data in Table 4, the parameters of this IOT experimental platform are as follows: the transmitting and receiving energy consumption is 50 nJ/bit, the number of nodes is less than 100, the initial node energy is 0.7 J, the time slice is 32 s, the voltage is 3.6 V, the data queue length is 15 kb, the transmission distance is greater than 20 m, the
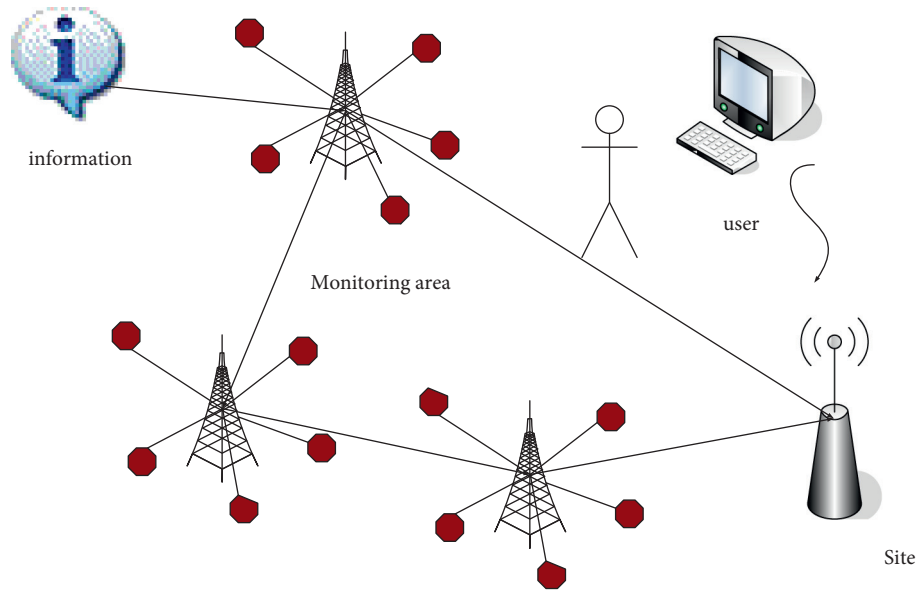
Figure 5: Structure diagram of wireless sensor network.

Table 1: System parameters.

Systems	Memory (GB)	Hard disk (GB)	Mainframe
Domain controller	3	30	2 core
Attendance system operating environment	3	30	1 core
Vcenter server	6	30	2 cores
Database server	3	30	2 cores
ESXI server	3	30	2 cores

Table 2: Experimental environment configuration table.

Configuration items		Parameters
Servers	64 GB	64 GB
	2 Gbps	2 Gbps
	4 cores	4 cores
	2 TB	2 TB
	64 bit	64 bit
Storage	16 TB	16 TB

Table 3: Comparison of detection algorithms.

Algorithms	Network	Parameters (MB)	mAP
RefineDet	VGG-16	138	82
Faster-RCNN	Resnet-50	122	77
EHN-B0	HourglassNet-B0	14	76
EHN-B3	HourglassNet-B3	36	83
StairNet	VGG-16	135	80

data processing energy consumption is 7 nJ/bit, the node data transmission rate is 130 kb/s, and the area is 200 m * 200 m.

According to the data in Figure 6, during the transition from 1 to 256 parallel output degrees, the normalized area of the computing unit is getting higher and higher. When the parallel output degree is 1, the normalized area is 0.35, and

when the parallel output degree is 2, the normalized area is 0.36. When the parallel output degree is 4, the normalized area is 0.37, when the parallel output degree is 8, the normalized area is 0.42, and when the parallel output degree is 16, the normalized area is 0.45. When the parallel output degree is 32, the normalized area is 0.5, when the parallel output degree is 64, the normalized area is 0.57, and when the parallel output degree is 128, the normalized area is 0.7. When the parallel output degree is 256, the normalized area is 0.98. These data show that the shared analog-to-digital conversion circuit can indeed effectively reduce the area of the computing array.

5. New Developments in Artificial Intelligence Algorithms for Big Data Processing of IOT Systems on Intelligent Production Lines

5.1. Data Reading Performance Analysis. In the actual use process, data sources may be diverse, there are multiple copies in each metadata, and any copy can send a metadata read request in response to the request of the client.

According to the data in Figure 7, four different copies were used for data reading and analysis in this experiment. During the reading process, the client is in an increasing state. According to the specific data, the data processing of Namenode copy is very general. When the number of clients

TABLE 4: Simulation environment parameters.

Name	Value	Name	Value
Energy consumption for transmitting and receiving	50 nJ/bit	Data queue length	15 kb
Number of nodes	Less than 100	Transmission distance	Greater than 20 m
Initial energy of nodes	0.7 J	Data processing energy consumption	7 nJ/bit
Time slice length	32 s	Node data transfer rate	130 kb/s
Voltage	3.6 V	Area	200 m * 200 m
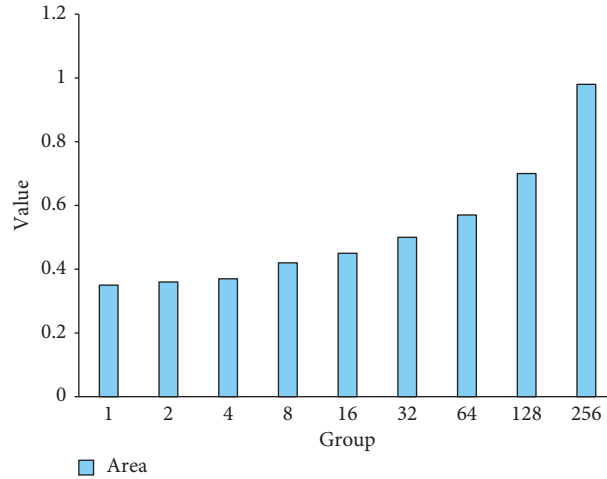
FIGURE 6: Output parallelism of computing unit.
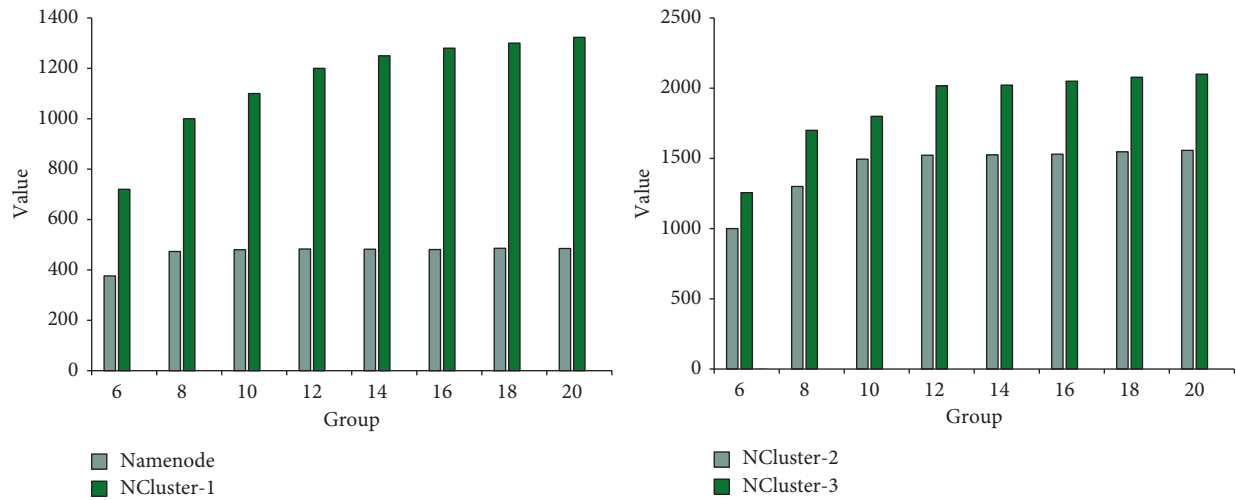
FIGURE 7: Data reading performance analysis.

is 6, 376 data are processed, when the number of clients is 8, 473 data are processed 473, and when the number of clients is 10, 480 data are processed. When the number of clients is 12, 483 data are processed, when the number of clients is 14, 482 data are processed, when the number of clients is 16, 481 data are processed, when the number of clients is 18, 486 data are processed, and when the number of clients is 20, 485 data are processed. According to this data, it can be seen that the Namenode mode reads data very poorly. When the number of clients reaches 8, the data processing basically

remains unchanged. When the NCluster-1 system is used for data reading, the data processing speed is improved, and as the number of clients increases, the reading speed becomes faster and faster. When the number of clients is 6, 720 data are processed. When the number is 20, 1323 data are processed, but according to specific data, when the number of clients reaches 12, the data reading speed of the NCluster-1 system will remain stable. The NCluster-2 system is used for data reading, when the number of clients is 6, 1000 data are processed, and when the number of clients is 20, 1557
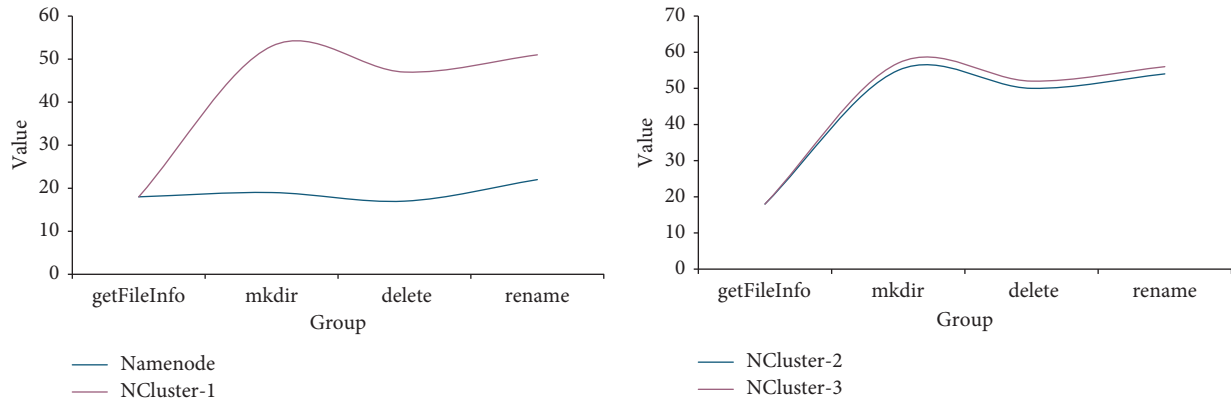
FIGURE 8: Data time-consuming comparative analysis.

data are processed. The NCluster-3 system is used for data reading, when the number of clients is 6, 1256 data are processed, when the number of clients is 20, 2100 data are processed, and the NCluster system will remain stable when the number of clients reaches 12.

According to the data in Figure 8, when the Namenode mode is used for data reading, the creation time is 18 ms, the running time is 19 ms, the deletion time is 17 ms, and the output result time is 22 ms. According to the data, the completion times of the four operations in the Namenode mode are relatively close. When using NCluster-1 mode for data reading, the creation time is 18 ms, the running time is 53 ms, the deletion time is 47 ms, and the output result time is 51 ms. When using NCluster-2 mode for data reading, the creation time is 18 ms, the running time is 55 ms, the deletion time is 50 ms, and the output result time is 54 ms. When using NCluster-3 mode for data reading, the creation time is 18 ms, the running time is 57 ms, the deletion time is 52 ms, and the output result time is 56 ms. According to the experimental data, when operating in NCluster mode, the running, deleting, and outputting times of the system have increased significantly, and there is no obvious difference between the creation time and the Namenode mode.

5.2. Big Data Processing Optimization Analysis. At present, big data is in a state of explosive growth, and big data has penetrated all aspects of daily life. For example, shopping websites that commonly used in daily life can recommend products to customers according to their preferences, saving time and improving user experience. In actual use, the performance of different modes is different.

According to the data in Figure 9, in the big data processing process, the data processing efficiency of different mapping methods is also different. The basic mapping of VGG1 is 20 Gop/s, and the hybrid mapping is 175 Gop/s. The basic mapping of VGG2 is 220 Gop/s, and the hybrid mapping is 590 Gop/s. The basic mapping of VGG3 is 430 Gop/s, and the hybrid mapping is 970 Gop/s. The basic mapping of VGG4 is 600 Gop/s, and the hybrid mapping is 793 Gop/s. The basic mapping of VGG5 is 1180 Gop/s, and the hybrid mapping is 1183 Gop/s. The basic mapping of LeNet1 is 5 Gop/s, and the hybrid mapping is 170 Gop/s. The

basic mapping of AlexeNet1 is 230 Gop/s, and the hybrid mapping is 600 Gop/s. According to the data, the performance of hybrid mapping in all modes is higher than that of basic mapping. The performance of VGG in the basic mode is 13 Gop/s, the mapping performance is 30 Gop/s, and the pipeline optimization performance is 231 Gop/s. The performance of AlexeNet in the basic mode is 38 Gop/s, the mapping performance is 40 Gop/s, and the pipeline optimization performance is 545 Gop/s. According to the data, the performance of the AlexeNet mode is better than that of the VGG mode.

5.3. Data Access Analysis. In the process of data access, the physical address and virtual address need to be converted, and there will be a large search loss during the operation. In order to determine the specific conversion situation, this article analyzes the different access modes. The details are as follows:

According to the data in Figure 10, the performance of the system is different in different operating modes. In automation mode, the system's thesaurus performance can be improved by 1 time, the index performance of the system has been improved by 1 time, the adjustment performance of the system has been improved by 1 time, the page performance of the system has been improved by 1 time, and the classification of the system performance has been doubled. In the DMA mode, the system's thesaurus performance can be improved by 1.65 times, the index performance of the system is improved by 1.59 times, the adjustment performance of the system is improved by 1.79 times, and the page performance of the system is improved by 2.23 times, and the system classification performance has been improved by 2.3 times. In atomic + cache mode, the system's thesaurus performance can be improved by 1.67 times, the system's exponential performance by 1.2 times, the system's adjustment performance by 1.6 times, the system's page performance by 1.5 times, and the classification performance has been improved by 1.9 times. In the atomic + DMA mode, the system's thesaurus performance can be improved by 2.35 times, the system's exponential performance by 2.41 times, the system's adjustment performance by 2.42 times, the system's page performance by 2.83 times, and the classification performance has been improved by 2.76 times.
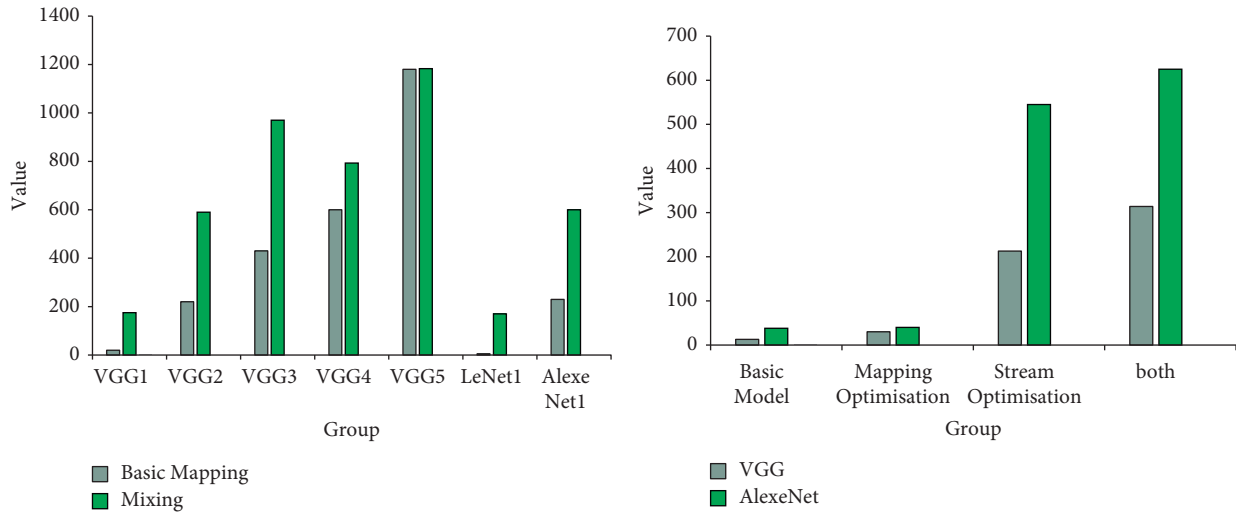
FIGURE 9: Big data processing optimization analysis.
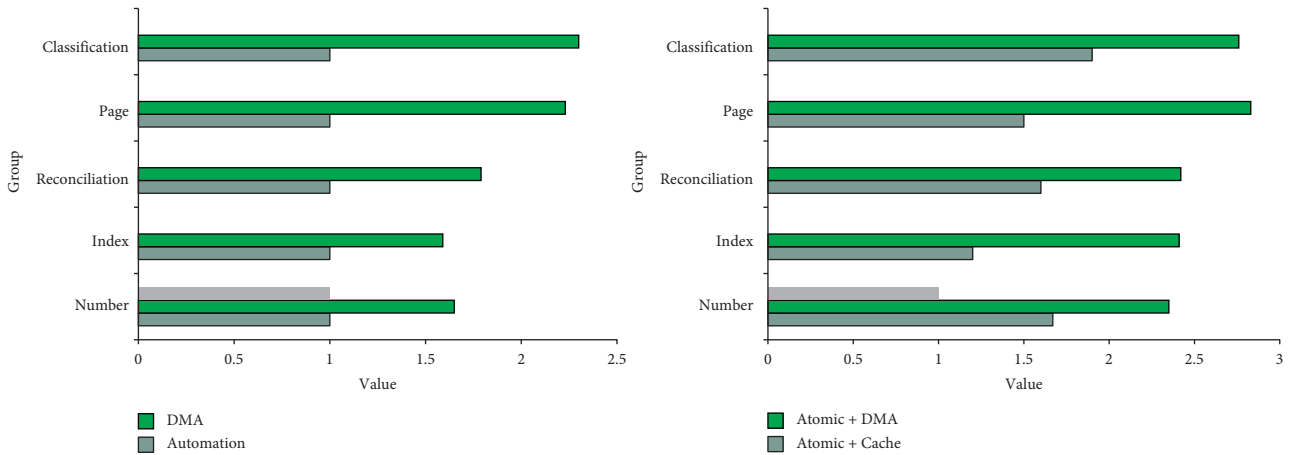
FIGURE 10: Analysis of data operation patterns.

6. Conclusions

The rapid development of China's economy has led to the development of science and technology, and artificial intelligence technology is the focus of current research. Artificial intelligence technology can help human beings enter into unknown areas to explore, expand the scope of knowledge, improve the speed and accuracy of human work, make full use of existing data, and complete the work better. However, due to the short development time of artificial intelligence technology, there are some shortcomings in the actual application process. Therefore, this article aims to study the new development of artificial intelligence algorithms for big data processing in the Internet of Things system on intelligent production lines. It is hoped that through the development of intelligent production lines and big data processing technologies, a way to optimize artificial intelligence algorithms can be found. Although some progress has been made, there are still some shortcomings: (1) this article aims to use the Internet of Things technology to deal with big data problems, but the data security issues are not considered in the actual application process. (2) The application field needs to be further broadened. Function optimization is only one aspect of the optimization field and solving different problems requires readjusting the model and parameters of the algorithm.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The author states that this article has no conflicts of interest.

Acknowledgments

References

[1] W. Wang, X. Zhu, L. Wang, Q. Qiu, and Q. Cao, "Ubiquitous robotic technology for smart manufacturing system," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 6018686, 14 pages, 2016.

[2] M. Liu, J. Ma, L. Lin, M. Ge, Q. Wang, and C. Liu, "Intelligent assembly system for mechanical products and key technology based on internet of things," *Journal of Intelligent Manufacturing*, vol. 28, no. 2, pp. 271–299, 2017.

[3] M. Wang, L. Zhao, R. Du et al., "A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms," *Applied Energy*, vol. 220, no. JUN.15, pp. 480–495, 2018.

[4] T Bao, L Cheng, B Chen et al., "Software design of reactive power optimization analysis based on artificial intelligence algorithms[J]," *Dianli Xitong Baohu yu Kongzhi/Power System Protection and Control*, vol. 46, no. 3, pp. 89–96, 2018.

[5] R. Dou and G. Nan, "Optimizing sensor network coverage and regional connectivity in industrial IoT systems," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1351–1360, 2017.

[6] M. N. Aman, K. C. Chua, and B. Sikdar, "Mutual authentication in IoT systems using physical unclonable functions," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1327–1340, 2017.

[7] T. Kim, D. M. Kim, N. Pratas, P. Popovski, and D. K. Sung, "An enhanced access reservation protocol with a partial preamble transmission mechanism in NB-IoT systems," *IEEE Communications Letters*, vol. 21, no. 10, pp. 2270–2273, 2017.

[8] S Yang, J Han, X Wang et al., "An obfuscatable Aggregatable signcryption scheme for unattended devices in IOT systems[J]," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 1067–1081, 2017.

[9] S Hu, "Research on data acquisition algorithms based on image processing and artificial intelligence[J]," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 06, pp. 1–13, 2020.

[10] B Gdla, C Ycl, Z C Wei et al., "A brief review of artificial intelligence applications and algorithms for psychiatric disorders[J]," *Engineering*, vol. 6, no. 4, pp. 462–467, 2020.

[11] A Me, A Af, B Mf et al., "Hybridization of artificial intelligence models with nature inspired optimization algorithms for lake water level prediction and uncertainty analysis - ScienceDirect," *Alexandria Engineering Journal*, vol. 60, no. 2, pp. 2193–2208, 2021.

[12] C, A Klein et al., "Intelligent furnace technology in continuous annealing and galvanizing lines," *Metallurgical Plant and Technology International*, pp. 42–47, 2018.

[13] I. Rodriguez, K. Nottensteiner, D. Leidner, M. Kassecker, F. Stulp, and A. Albu-Schaffer, "Iteratively refined feasibility checks in robotic assembly sequence planning," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1416–1423, 2019.

[14] Y. Wang, S. Men, and T. Guo, "Application of blockchain technology in value chain of procurement in manufacturing enterprises," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1674412, 8 pages, 2021.

[15] R. Allen and D. Masters, "Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making," *ERA Forum*, vol. 20, no. 4, pp. 585–598, 2020.

[16] P. Lade, R. Ghosh, and S. Srinivasan, "Manufacturing analytics and industrial internet of things," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 74–79, 2017.

[17] A H Azimi and A A Abdoos, "A new intelligent method for parallel transmission lines protection," *Technology and Economics of Smart Grids and Sustainable Energy*, vol. 5, no. 1, pp. 1–11, 2020.

[18] A. Schmidt, F. Schellroth, M. Fischer, L. Allimant, and O. Riedel, "Reinforcement learning methods based on GPU accelerated industrial control hardware," *Neural Computing and Applications*, vol. 33, no. 18, pp. 12191–12207, 2021.

[19] H. Fazlollahtabar and S. T. Akhavan Niaki, "Integration of fault tree analysis, reliability block diagram and hazard decision tree for industrial robot reliability evaluation," *Industrial Robot: International Journal*, vol. 44, no. 6, pp. 754–764, 2017.

[20] N. Nahas, "Buffer allocation and preventive maintenance optimization in unreliable production lines," *Journal of Intelligent Manufacturing*, vol. 28, no. 1, pp. 85–93, 2017.

[21] V. T. Yen, W. Y. Nan, and P. Van Cuong, "Recurrent fuzzy wavelet neural networks based on robust adaptive sliding mode control for industrial robot manipulators," *Neural Computing and Applications*, vol. 31, no. 11, pp. 6945–6958, 2019.

[22] Z. Yushu, H. Qi, C. Guo, Z. Xinpeng, and X. Yong, ""A low-overhead, confidentiality-assured, and authenticated data acquisition framework for IoT"," *IEEE Transactions on Industrial Informatics, Publication Year*, p. 1, 2019.

[23] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, 2016.

[24] F. Xiao, "Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy," *Information Fusion*, vol. 46, pp. 23–32, 2019.

[25] I. K. Osamh and G. M. Abdulsahib, "Energy efficient routing and reliable data transmission protocol in WSN," *International Journal of Advances in Soft Computing and Its Applications*, vol. 12, no. 3, pp. 45–53, 2020.

[26] X. Li, H. Jianmin, B. Hou, and P. Zhang, "Exploring the innovation modes and evolution of the cloud-based service using the activity theory on the basis of big data," *Cluster Computing*, vol. 21, no. 1, pp. 907–922, 2018.

[27] S. Wan, X. Li, Y. Xue et al., "Efficient computation offloading for Internet of Vehicles in edge computing-assisted 5G networks," *The Journal of Supercomputing*, vol. 76, 2019.

Research Article

A Tropical Cyclone Center Location Method Based on Satellite Image

Qingxiang You [ID],[1,2] Zhenqing Li [ID],[2] Cheng Qian [ID],[1] and Tian Wang [ID][1]

[1]*Department of Computer and Information Engineering, Changzhou Institute of Technology, Changzhou, China*
[2]*University of Shanghai for Science and Technology, Shanghai, China*

Correspondence should be addressed to Qingxiang You; youqx@czu.cn

Accurately detecting and locating the center of the tropical cyclone is critical for the trajectory forecasting. This study proposed an automatic method for centers' location of the tropical cyclones based on the visible or the infrared satellite images. The morphological structure of the tropical cyclone is modeled using the circular pattern. The tropical cyclone center is located based on regional pixels instead of skeleton points. All pixels in a segmented cloud cluster vote for a 2-dimensional accumulator. The center of the cloud cluster is computed by the mean voting distances, which are calculated by fitting quadratic functions in every column of the two-dimensional (2D) accumulator. Then, a linear function is fitted according to the functional relationship between the mean voting distance and voting angle. The fitted coefficients of the linear function are the center coordinates of the tropical cyclone. The proposed method for centers location of the tropical cyclones is tested using visible and infrared satellite images. The results of center location are compared with the best track provided in JMA datasets.

1. Introduction

The tropical cyclone is a kind of atmospheric circulation systems, which may cause disasters and lead to economic loss in coastal areas. The trajectory forecast [1, 2] of tropical cyclones is very significant in order to avoid the destruction by tropical cyclones. Meteorological satellites offer a reliable solution to observe tropical cyclones [3, 4]. The trajectory of a tropical cyclone can be analyzed and estimated based on satellite images [5]. The first comprehensive technique for analyzing tropical cyclones using satellite images was Dvorak techniques [6].

The center location for a tropical cyclone is important for the trajectory forecast of the tropical cyclone [7–9]. The center position of the tropical cyclone is estimated by finding the spiral origin [10] or by fitting the elliptic center [11]. Huadong et al. [12] and Ryglicki and Hart [13] investigated various methods of tropical cyclone center location and classified methods into two categories: morphological pattern analysis method and wind field analysis method.

The first kind of method is based on pattern analysis of morphological characteristics of the tropical cyclone [14, 15]. By using image processing techniques [16], the spiral center of the cyclone is determined by finding the origin of a logarithmic helix [17] or by locating the point where the gradient vectors of brightness temperature are converging [18]. In order to detect the tropical cyclone center, the density matrix [19] and the deviation angle variance [20] are exploited. The located center is accurate, but it does not work well when the morphological structure of the tropical cyclone is not apparent.

The second kind of method is based on the wind field analysis [21, 22]. The center of tropical cyclone is determined by finding the minimum value of wind speed or the maximum value of cyclone vortices [23]. In order to locate the center position of the tropical cyclone, a mathematical morphology method [24] or saliency detection method [25] are utilized. It can deal with the center location of weak circulation, but the wind field inversion is affected.

The current methods for center location of the tropical cyclone require applying an edge detector to extract skeleton

points in satellite cloud images. However, the extracted skeleton is usually inaccurate due to noise and disturbance. The tropical cyclones are often segmented from the images before the circle detection, so the shape of the tropical cyclone relies on the segmentation result. When the noise is incorporated into the segmentation result, the shape of the tropical cyclone is distorted. The circular skeleton is not enough definite for the successive circle detection. It will lead to the failure in the circle detection. In order to overcome the problem of edge extraction, the tropical cyclone center is analyzed and located based on regional pixels instead of skeleton points. The methods based on Hough transform can be used to detect circles [26, 27]. The voting of parameters determines the locations and the sizes of the circles. Using a modified Hough transform [28], the center position of the tropical cyclone is determined by fitting a linear function according to the functional relationship of the mean voting distance, which is calculated in each column of the accumulator by fitting a quadratic function.

2. Methods

The satellite cloud image is segmented; firstly, the segmented cloud cluster with the most pixels is selected. After regional pixels' vote, the center of a cloud cluster is determined by function fitting techniques. The proposed algorithm is made up of four steps: image binarization, Hough vote, searching voting distances, and solving the center coordinates. The flowchart of the proposed algorithm is shown in Figure 1.

2.1. Satellite Cloud Image Segmentation. The image is segmented using seed expanding and the threshold constraint [29–31]. Two thresholds are defined. The first threshold T1 is defined as the quantity of pixels whose intensity greater than T1 is less than 300. The second threshold T2 is defined as the quantity of pixels whose intensity greater than T2 is less than 30000. A pixel, whose intensity is greater than T1, can be selected as a seed. If the intensity of its neighboring pixel is greater than T2, the region is expanded [32].

The region with the most pixels is considered as a candidate of the tropical cyclone. An example of cloud image segmentation is shown in Figure 2. Although the shape of a segmented cloud cluster is irregular, it is modeled using a near-circular pattern. The circle center is considered as the center of a cloud cluster corresponding to a tropical cyclone.

2.2. Regional Pixels' Voting. In order to extract the center of the near-circular cloud cluster, a modified Hough transform is used. Instead of skeleton points, the regional pixels vote for a 2-dimensional accumulator. The voting formula is

$$\rho = x \cdot \cos \theta + y \cdot \sin \theta, \quad \theta \in [0 \ \pi), \tag{1}$$

where θ is the Hough voting angle and ρ is the computed voting distance. For a voting angle θ_i, the validated voting distance ρ is limited. It is shown in Figure 3. ρ_m is the mean voting distance with respect to each voting angle θ. It is also the regional center voting for the 2D accumulator. The voting value $H(\theta, \rho)$ increases from 0 to the maximum and
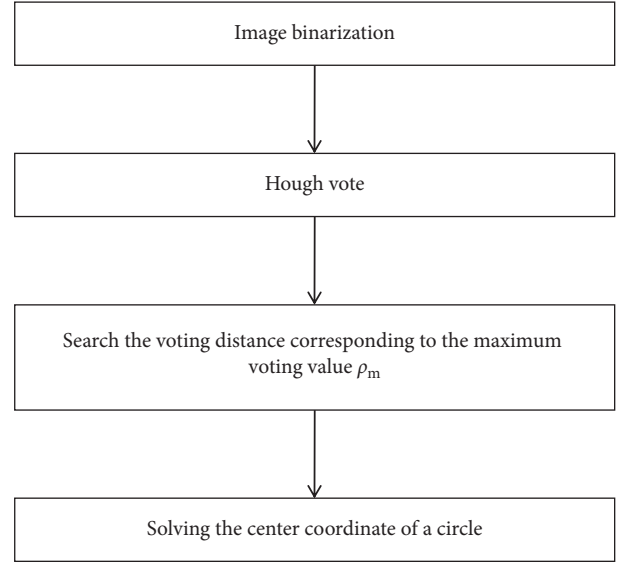
FIGURE 1: The flowchart of the proposed algorithm.
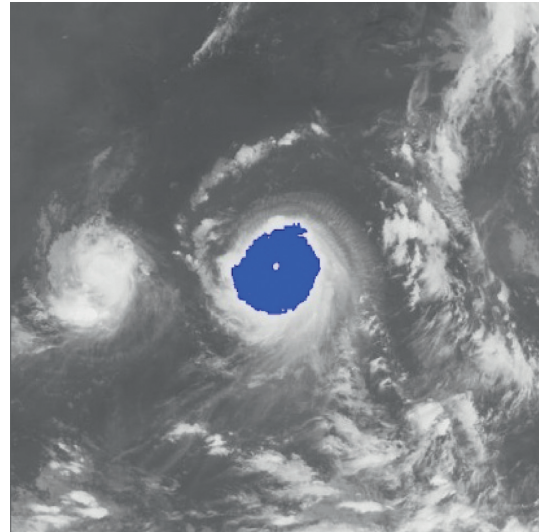
FIGURE 2: Extraction of the cloud cluster.

then decreases to 0. Let $\rho_m(\theta)$ be the mean voting distance corresponding to the voting angle θ.

After all regional pixels have voted for the 2D accumulator, a quadratic function is fitted using the voting information in each column of the 2D accumulator.

2.3. Quadratic Functions' Fitting. The voting value $H(\theta, \rho)$ corresponding to distance ρ_j and angle θ_i is illustrated in Figure 4. The voting value is

$$H(\rho, \theta) = 2 \cdot \sqrt{r^2 - (\rho - \rho_m)^2}. \tag{2}$$

Thus, the functional relationship of the voting value $H^2(\theta, \rho)$ with respect to the voting distance ρ is a quadratic polynomial function, whose maximum is located at the mean voting distance ρ_m. Therefore, ρ_m can be obtained by
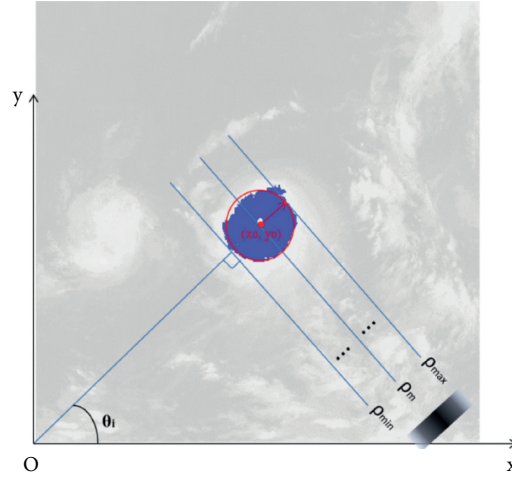
FIGURE 3: Voting of regional pixels.
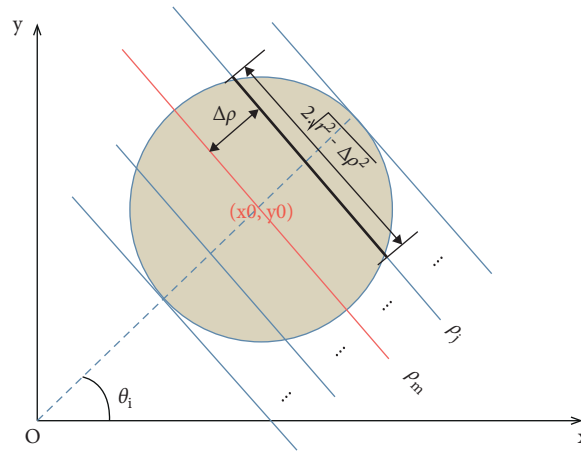
FIGURE 4: Voting value $H(\theta,\rho)$.

fitting a quadratic function using the voting values in each column of the 2D accumulator.

In column θ_i of the 2D accumulator, all cells with nonzero votes are searched, and a quadratic function f is fitted to the data pairs (H^2, ρ). The fitted function is denoted as

$$f: H^2(\rho, \theta) = a_2\rho^2 + a_1\rho + a_0. \tag{3}$$

Figure 5 shows an example of the quadratic function fitting. The mean voting distance is determined by

$$\rho_m = \rho \left| \frac{\partial f}{\partial \rho} = 0. \right. \tag{4}$$

2.4. Linear Function Fitting. For each voting angle θ, the computed mean voting distance $\rho_m(\theta)$ is located at ρ whose voting value is the maximum. Following Figures 3 and 4, the line corresponding to maximum voting value passes through the center of the circular region, which means the center

(x_0, y_0) of a cloud cluster always votes for the distance ρ_m; therefore, there is the following formula:

$$\rho_m(\theta) = x_0 \cdot \cos\theta + y_0 \cdot \sin\theta. \tag{5}$$

Thus, the relationship between the computed mean voting distance $\rho_m(\theta)$ and the voting angle θ is a sine function, whose coefficients rely on the coordinates (x_0, y_0) of the region center. The center can be calculated by fitting a sine function. In order to calculate the center coordinates, conveniently, the sine function is linearized as

$$\frac{\rho_m(\theta)}{\cos\theta} = x_0 + y_0 \cdot \tan\theta. \tag{6}$$

Thus, the functional relationship between $\rho_m(\theta)/\cos(\theta)$ and $\tan(\theta)$ is linear. We fit a linear function g using data pairs $(\rho_m(\theta_i)/\cos(\theta_i), \tan(\theta_i))$. When the voting angles equal 40°, 48°, 56°, 64°, 72°, 80°, 88°, and 96°, the corresponding data pairs $(\rho_m(\theta_i)/\cos(\theta_i), \tan(\theta_i))$ are shown in Table 1, and the linear function fitting is illustrated as Figure 6. In addition, to avoid the setting of θ to 90°, θ is often sampled at the discrete values around 90°.

$$H^2(\theta, \rho) = -2.6156\rho^2 + 1874\rho - 330374$$
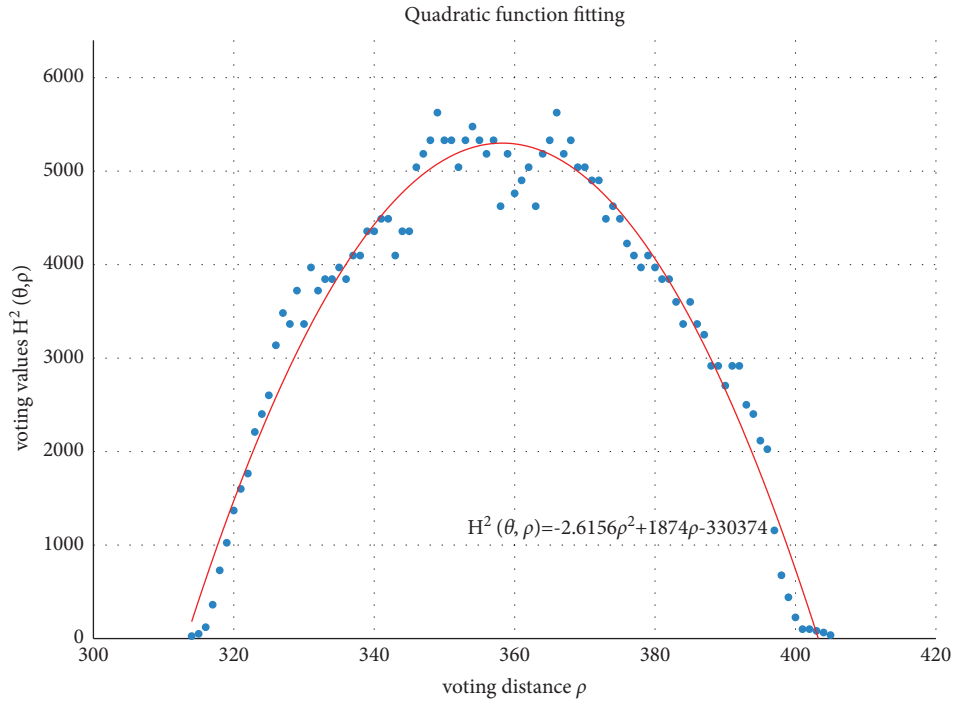
Figure 5: Quadratic function fitting.

Table 1: The relationship between $\rho_m(\theta)/\cos(\theta)$ and $\tan(\theta)$.

θ	40	48	56	64	72	80	88	96
ρ_m	358.2352042	350.6212455	336.8496554	316.8497009	290.6307249	259.0087804	259.0087804	259.0087804
$\tan(\theta)$	0.839099631	1.110612515	1.482560969	2.050303842	3.077683537	5.67128182	28.63625328	−9.514364454
$\rho_m/\cos(\theta)$	467.6428468	523.9952292	602.385426	722.7886763	940.5007823	1491.572119	7421.562053	−2477.878008

$$y = 212.04x + 288.53$$
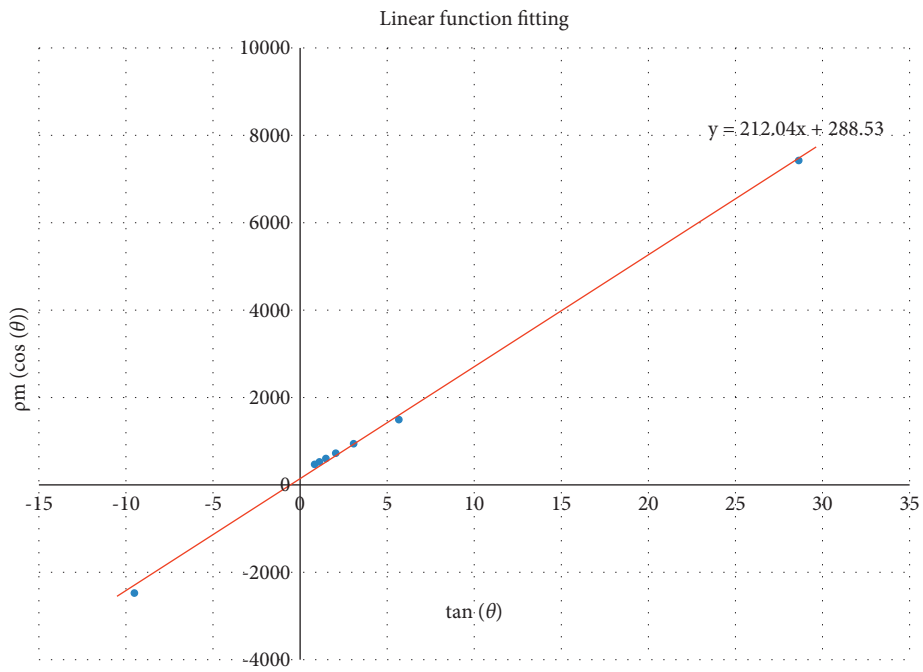
Figure 6: Linear function fitting.

The center coordinates (x_0, y_0) of the cloud cluster happen to be the fitted coefficients. The calculated center of the tropical cyclone is labeled in Figure 7.

3. Results

3.1. Implementation Details. The dataset used to test the proposed method is from Meteorological Satellite Observation Images of Typhoon Saomai (No. 200608) and Maria (No. 201808) provided by the Japan Meteorological Agency (JMA). Both the visible images and the infrared images are tested. With the satellite observation images, the cloud cluster of the tropical cyclone is extracted and its center is located.

The satellite cloud images are segmented using seed expanding techniques and threshold constraints. The segmented cloud cluster with the most pixels is recognized as the tropical cyclone; its center is calculated by voting and fitting. The center location results on two visible images and two infrared images are shown.

3.2. Experiments' Analysis. In Figure 8, it is noted that, despite of either the visible images or the infrared images, the proposed method is able to extract the regions of tropical cyclone. Based on the regions, the centers of the tropical cyclone can be further determined. In addition, the method can also deal with both eye cyclones and noneye cyclones. It is largely attributed to the fitting for the shape of the cyclone regions. Even though the shapes of tropical cyclones are irregular, the shapes are still approximated by circles. Once the circles are found, the centers can be determined. The located centers of typhoons are also compared with the Best Track (BT), and the averaged location errors are shown in Table 2. Because the size of typhoons is in itself large enough, these location errors can be tolerated. The experimental results show that the proposed center location method provides reliable estimates of tropical cyclone centers.

As illustrated in Figures 9 and 10, the segmentation of the typhoons region has great influence on the determination of the centers. When the segmented region covers most of the tropical cyclone, the shape of the cloud region is more likely to be a circle. As a result, the accuracy in the fitting of the circle is significantly improved. In Figure 7(a), while the tailor of the tropical cyclone is discarded, the segmentation maintains the central part of the tropical cyclone. The central part almost covers the whole tropical cyclone. Compared with the segmented cloud in Figure 9(a), the tropical cyclone in Figure 10(a) undergoes the deformation. It is obvious that the shape of the tropical cyclone in Figure 10(a) is irregular. It further causes the error in the estimation of the center of the tropical cyclone. Although it is inevitable for the error to be produced, the increase in the number of voting angle θ will reduce the error.

Like other methods for centers' location, the estimation of centers depends on the segmentation of the typhoons' region. The locating results are improved by using function techniques. A set of quadratic functions are fitted to voting values corresponding voting distances, and a linear function

x₀=212.04   y₀=288.53
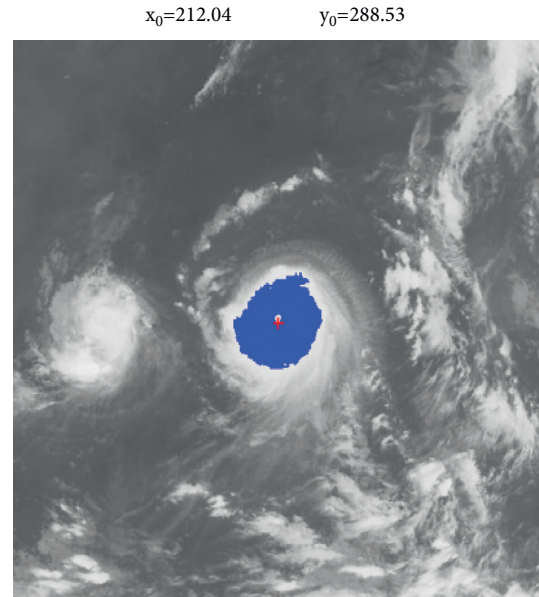
FIGURE 7: Center location.

is fitted to mean voting distance with respect to the voting angle.

4. Discussion

In general, due to the existence of cloud tailors, small cloud clusters, and so on, it is difficult for the segmentation to generate the result of the tropical cyclone with the regular circular shape. Hence, it is important for the method to resist against the noises inherent to the segmentation results. The proposed method alleviates the errors in the estimation of centers caused by the irregular typhoons regions. Similar to other methods based on the morphological pattern analysis, the proposed method also relies on the organization structure of the tropical cyclone. Different with other methods, the proposed method uses not only skeleton points but also regional pixels. Therefore, the edge extraction, which is very sensitive to disturbance and noise, is not necessary. The proposed method is more robust than methods based on skeleton information. The robustness is largely attributed to the setting of multiple voting angles.

From Figures 11–18, the voting process of the parameters related to the circular region covering the typhoons region in Figure 2 is shown. At the same time, the fitting process is also given in these figures. In detail, a quadratic function is adopted to fit the relationship between the voting angle and the voting distance. To acquire the two-dimensional center location, it requires that at least two fitting functions are offered. In fact, the typhoons' region is often irregular, so the center location solved by two fitting functions is coarse. It means that more fitting functions are needed so as to give rise to an accurate estimation. In Figures 11–18, the voting angles are, respectively, sampled at 40°, 48°, 56°, 64°, 72°, 80°, 88°, and 96°. As a result, eight fitting functions are obtained. Among them, while the estimations of the voting distance in Figures 11 and 12 nearly distribute

(a)

(b)
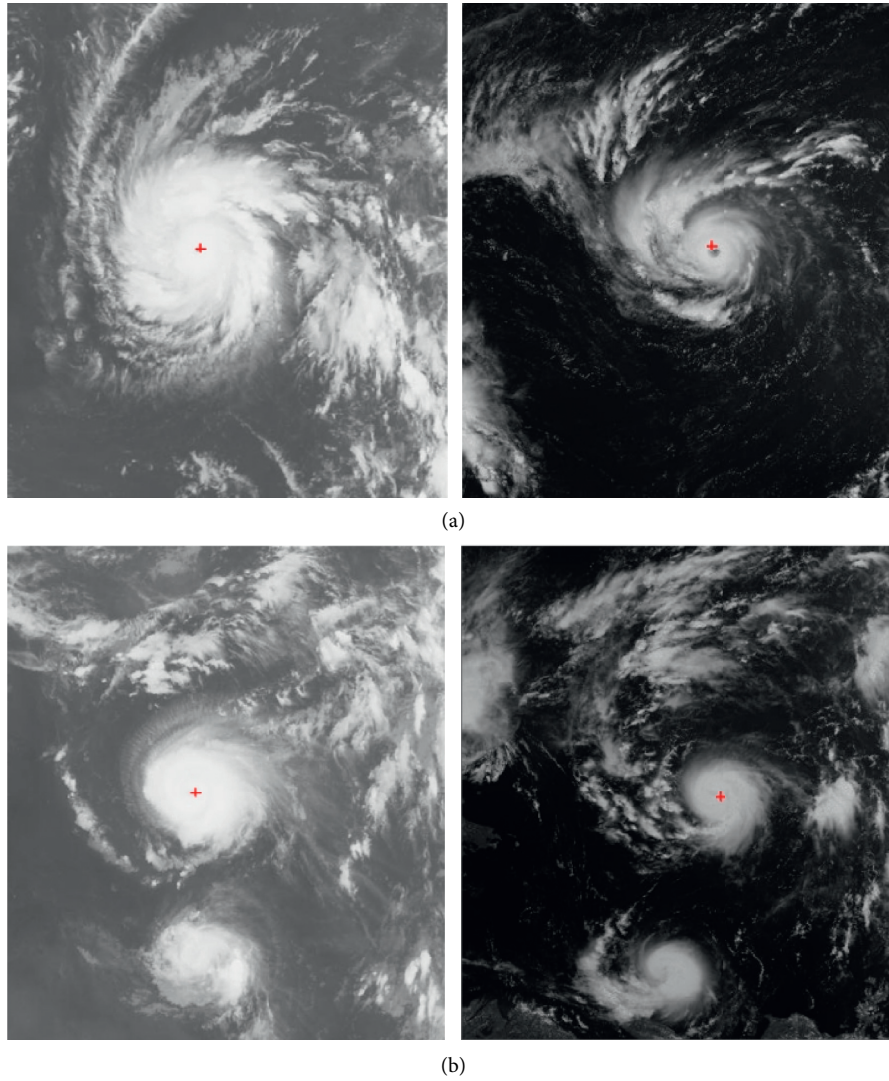
Figure 8: Center estimates for eye and noneye cyclones in visible and infrared satellite images. (a) Satellite images in the infrared channel. (b) Satellite images in the visible channel.

Table 2: The averaged errors of center location.

Tropical cyclones	Saomai (km)	Maria (km)
Averaged errors	18.62	20.56

(a)   (b)
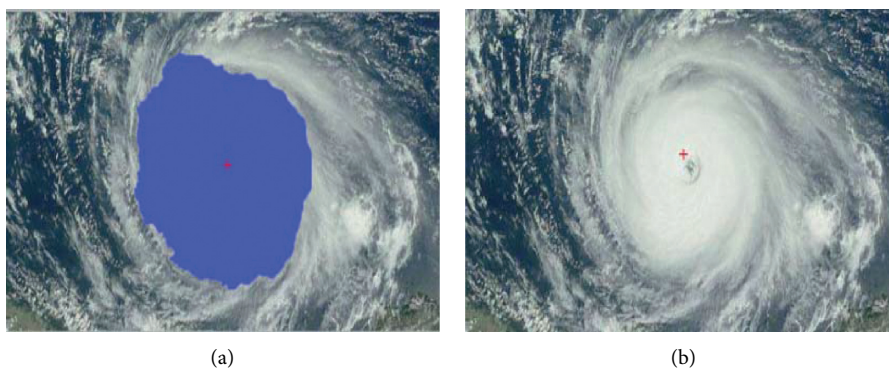
Figure 9: (a) The region of the tropical cyclone is extracted from a visible satellite image by the segmentation. It is filled with the blue color. (b) The center of the tropical cyclone is determined based on the region of the tropical cyclone.

(a)   (b)
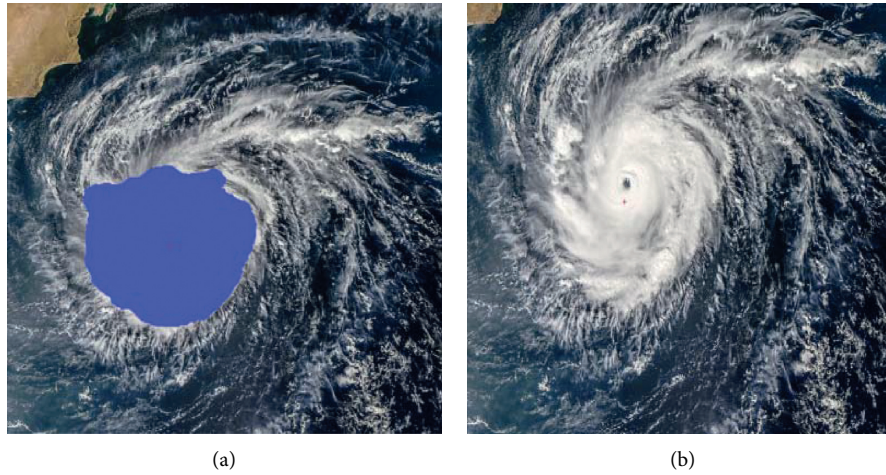
FIGURE 10: (a) The region of the tropical cyclone is extracted from a visible satellite image by the segmentation. It is filled with the blue color. (b) The center of the tropical cyclone is determined based on the region of the tropical cyclone.

Quadratic function fitting (θ=40)

$H^2(\theta, \rho)$=-2.6156ρ^2+1874ρ-330374
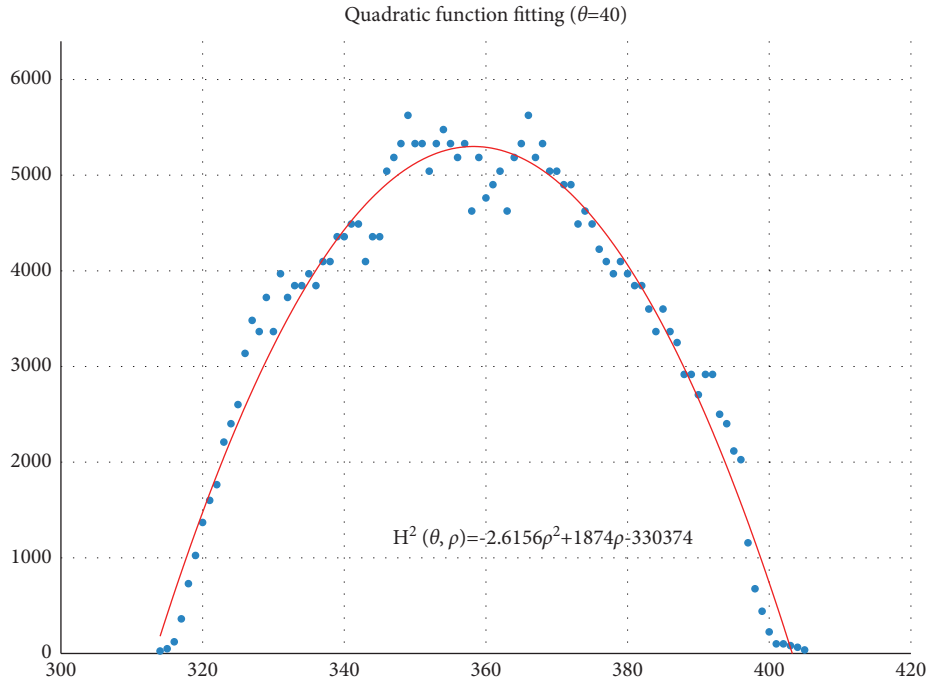
FIGURE 11: The voting angle θ is set to 40°. When the voting distance ρ is 358.24, $H(\theta,\rho)$ reaches the peak.

around the quadratic fitting functions, some points in Figures 13 and 14 significantly deviate from the fitting function. Along the lines with the angles 56° and 64°, the invagination of the shape inhibits the points from scattering around the quadratic functions. These points are more likely to be outliers. Hence, if only the points in Figures 13 and 14 are given, the quadratic functions estimated by these points cannot accurately represent the boundary enclosing the tropical cyclone. In Figures 17 and 18, θ is sampled at 88° and 96°, respectively, which are close to 90°. It avoids the setting of θ to the trivial value of 90°.

In the proposed method, the center location is usually located at the line with maximal voting distance. However,

once the shape of the typhoon region is an irregular circle, it is possible for the estimated center to be far away from the real center of tropical cyclone. The usage of the fitting function removes the lines existing in the form of outliers. In Figure 13, the distribution of the points representing the voting distances in the proximity of the peak of the quadratic function is in chaos. Since that, in theory, the voting values corresponding to the regular circle should satisfy the quadratic function, the quadratic fitting to the function excludes those values that are not in favor of the estimation of the circle. Hence, it is obvious that the fitting to the quadratic function enhances the robustness of the estimation. The fitting functions with respect to more voting angles lead to
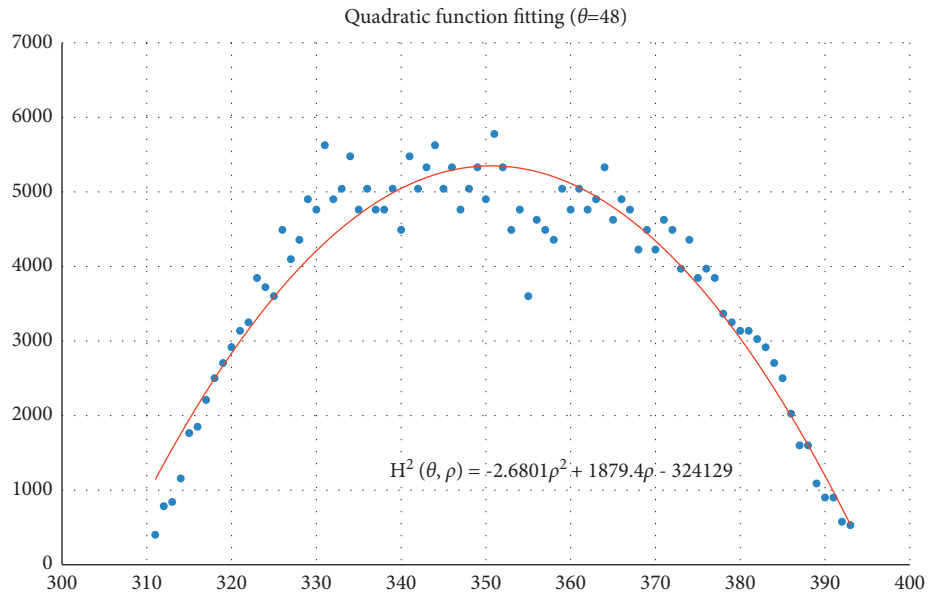
FIGURE 12: The voting angle θ is set to 48°. When the voting distance ρ is 350.62, $H(\theta,\rho)$ reaches the peak.

In the chart: Quadratic function fitting (θ=48)

$H^2(\theta, \rho) = -2.6801\rho^2 + 1879.4\rho - 324129$
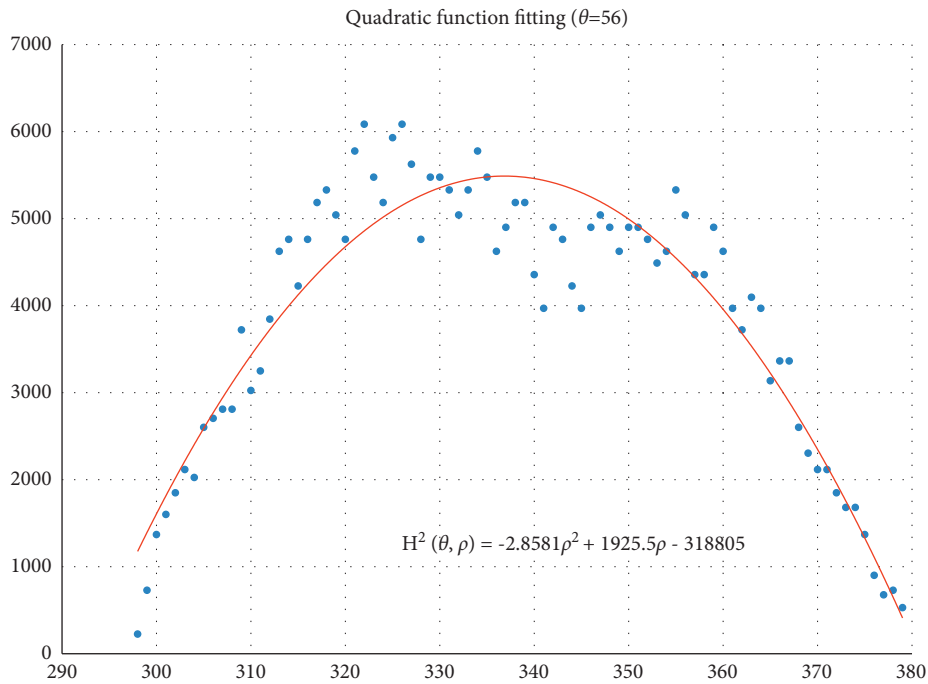
FIGURE 13: The voting angle θ is set to 56°. When the voting distance ρ is 336.85, $H(\theta,\rho)$ reaches the peak.

In the chart: Quadratic function fitting (θ=56)

$H^2(\theta, \rho) = -2.8581\rho^2 + 1925.5\rho - 318805$

more candidate centers. The regression over these centers can generate the reliable center. However, more fitting functions incur heavier computation burden, and it reduces the speed of finding the centers of tropical cyclones. Sometimes, the speed is important for the alert of the tropical cyclones.

In addition, the method is not sensitive to the type of images. Whatever the infrared image or the visible image is, the cloud cluster can be approximated as a near-circular region. Therefore, the center location method can be applied to the visible channel and infrared channel of meteorological satellite images.

Quadratic function fitting (θ=64)

$$H^2(\theta, \rho) = -3.1597\rho^2 + 2002.3\rho - 311499$$
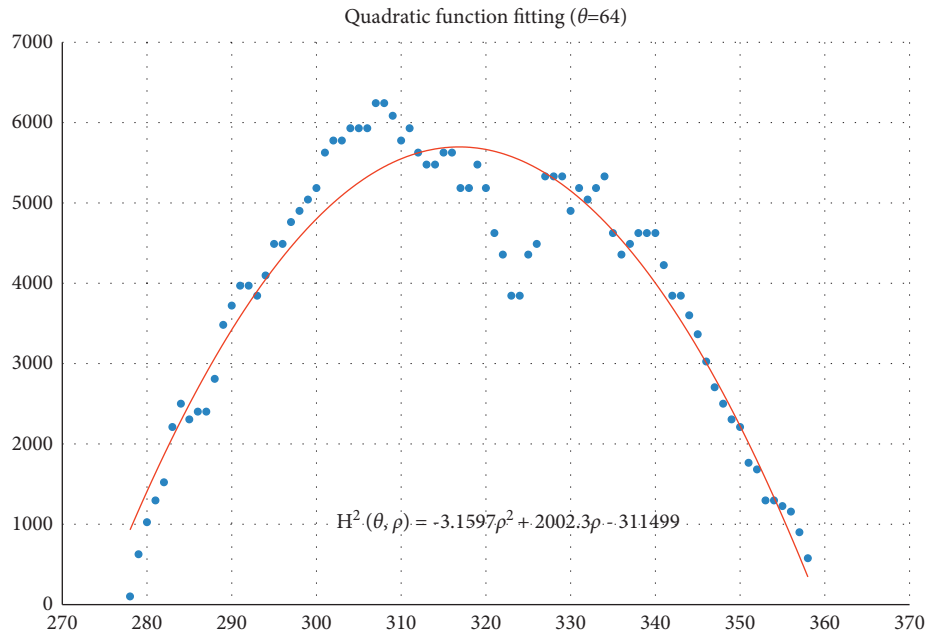
FIGURE 14: The voting angle θ is set to 64°. When the voting distance ρ is 316.85, $H(\theta,\rho)$ reaches the peak.

Quadratic function fitting (θ=72)

$$H^2(\theta, \rho) = -3.4389\rho^2 + 1998.9\rho - 284553$$
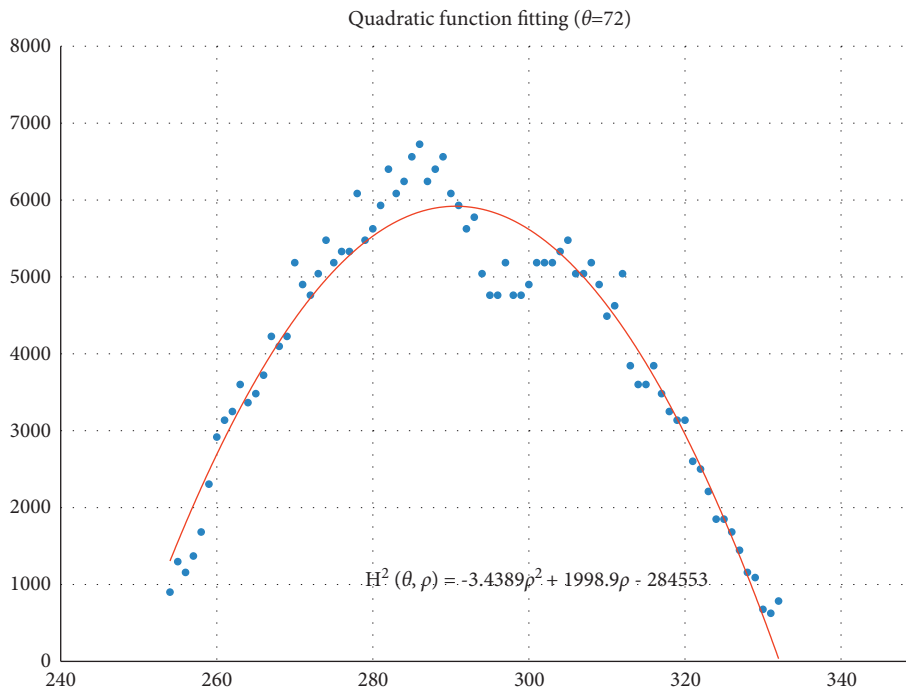
FIGURE 15: The voting angle θ is set to 72°. When the voting distance ρ is 290.63, $H(\theta,\rho)$ reaches the peak.

Quadratic function fitting (θ=80)

$$H^2(\theta, \rho) = -3.9406\rho^2 + 2041.3\rho - 258135$$
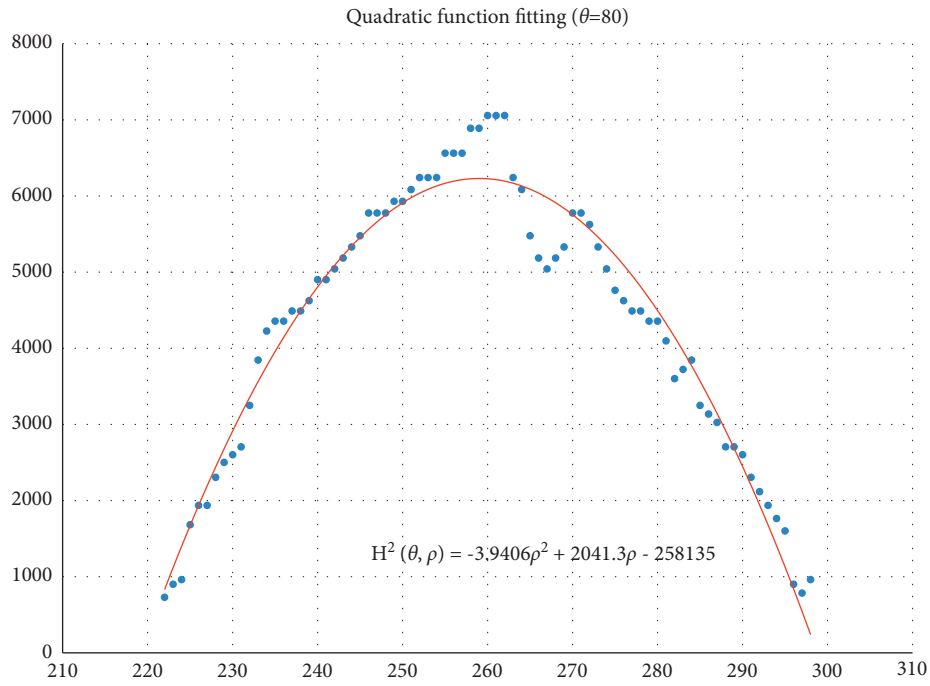
FIGURE 16: The voting angle θ is set to 80°. When the voting distance ρ is 259.01, $H(\theta,\rho)$ reaches the peak.

Quadratic function fitting (θ=88)

$$H^2(\theta, \rho) = -4.1954\rho^2 + 1868.6\rho - 201590$$
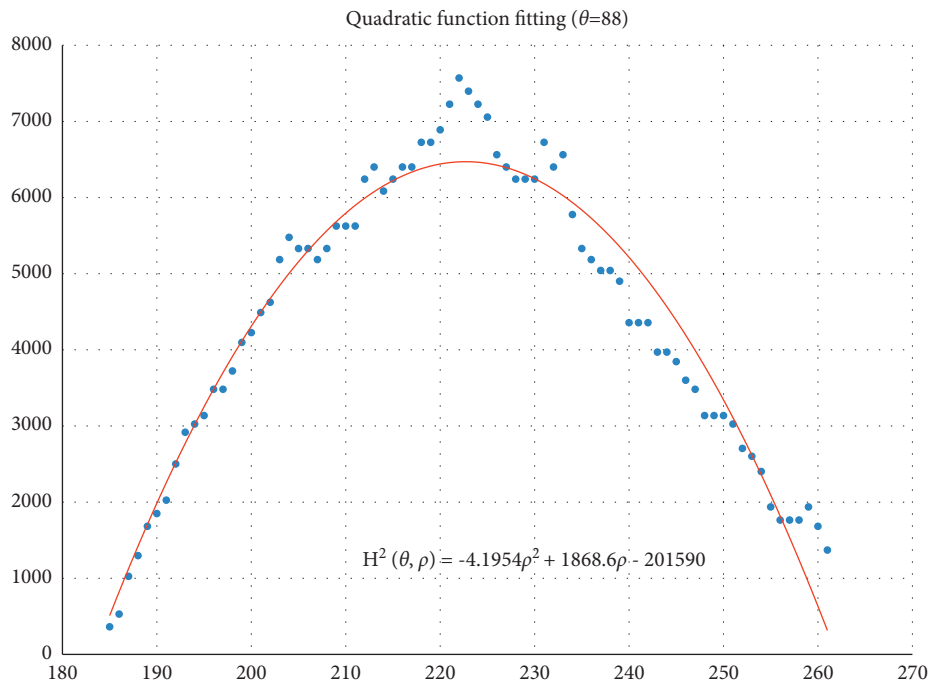
FIGURE 17: The voting angle θ is set to 88°. When the voting distance ρ is 259.01, $H(\theta,\rho)$ reaches the peak.
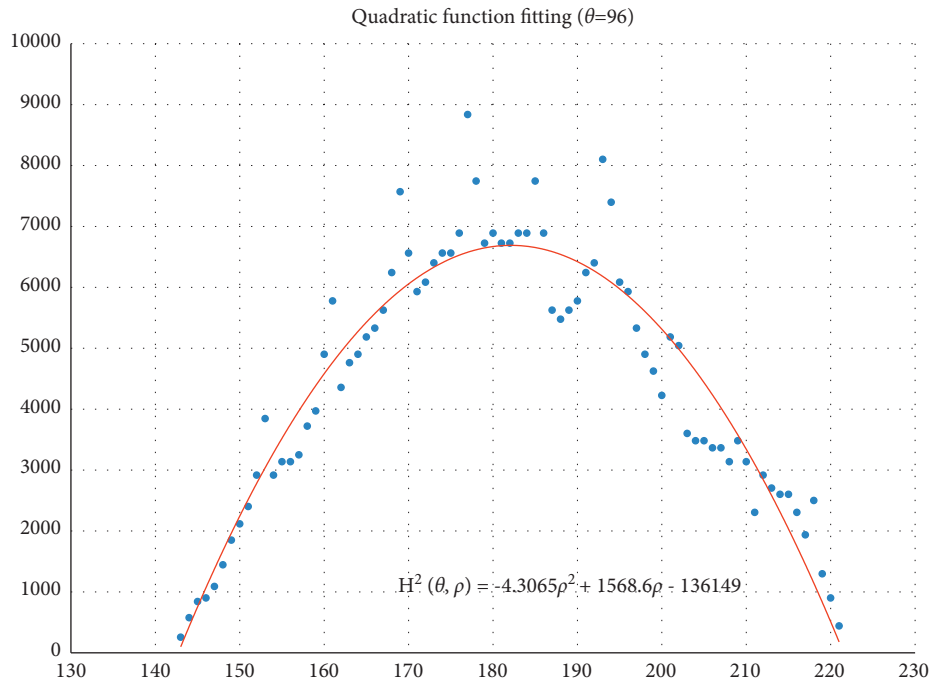
FIGURE 18: The voting angle θ is set to 96°. When the voting distance ρ is 259.01, $H(\theta,\rho)$ reaches the peak.

5. Conclusions

An automatic method for cyclone center location is proposed based on regional pixels instead of skeleton points. The morphological structure of the tropical cyclone is modeled using the circular pattern. The skeleton extraction is avoided. By fitting a linear function, the center coordinates of the tropical cyclone happen to be the coefficients of the fitted function. The method can deal with both visible images and infrared images and both eye cyclones and noneye cyclones. The reliable estimates of tropical cyclone centers are obtained in spite of the variability of TC morphological structure in visible or infrared images. At present, the thresholds for the image binarization are manually selected. It cannot adapt itself to the segmentation requirement. In the future, it is possible for the adaptive image binarization to be brought into the image segmentation.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence our work and there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled "A Tropical Cyclone Center Location Method Based on Satellite Image."

References

[1] R. Mario, L. Sangseung, J. Soohwan, and Y. Donghyun, "Prediction of a typhoon track using a generative adversarial network and satellite images," *Scientific Reports*, vol. 9, no. 1, p. 6057, 2019.

[2] F. P. Paul and D. B. Subrahamanyam, "Prediction of tropical cyclone trajectories over the Northern Indian Ocean using COSMO," *Meteorology and Atmospheric Physics*, vol. 133, pp. 1–14, 2021.

[3] H. Xiao and M. Wei, "An early warning method for sea typhoon detection based on remote sensing imagery," *Journal of Coastal Research*, vol. 82, pp. 200–205, 2018.

[4] C. Zhang, X. Wang, L. Ma, and X. Lu, "Tropical cyclone intensity classification and estimation using infrared satellite images with deep learning[J]," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, p. 1, 2021.

[5] T. L. Olander and C. S. Velden, "The advanced dvorak technique (adt) for estimating tropical cyclone intensity: update and new capabilities," *Weather and Forecasting*, vol. 34, no. 4, pp. 905–922, 2019.

[6] V. F. Dvorak, "Tropical cyclone intensity analysis and forecasting from satellite imagery," *Monthly Weather Review*, vol. 103, no. 5, pp. 420–430, 1975.

[7] P. Wang, P. Wang, C. Wang, Y. Yuan, and D. Wang, "A center location algorithm for tropical cyclone in satellite infrared

images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2161–2172, 2020.

[8] X. Lu, H. Yu, X. Yang, X. Li, and J. Tang, "A new technique for automatically locating the center of tropical cyclones with multi-band cloud imagery," *Frontiers of Earth Science*, vol. 13, no. 4, pp. 836–847, 2019.

[9] T. Hu, Y. Wu, G. Zheng, D. Zhang, Y. Zhang, and Y. Li, "Tropical cyclone center automatic determination model based on HY-2 and QuikSCAT wind vector products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1–13, 2019.

[10] B. S. Yurchak, "Estimation of tropical cyclone intensity from the satellite infrared images of its spiral cloud bands," *Russian Meteorology and Hydrology*, vol. 43, no. 9, pp. 581–590, 2018.

[11] S. Chaurasia, C. M. Kishtawal, and P. K. Pal, "An objective method of cyclone centre determination from geostationary satellite observations," *International Journal of Remote Sensing*, vol. 31, no. 9-10, pp. 2429–2440, 2010.

[12] Y. Huadong, W. Liguang, and X. Tong, "Comparisons of four methods for tropical cyclone center detection. In a high-resolution simulation," *Journal of the Meteorological Society of Japan*, vol. 98, no. 2, pp. 379–393, 2020.

[13] D. R. Ryglicki and R. E. Hart, "An investigation of center-finding techniques for tropical cyclones in mesoscale models," *Journal of Applied Meteorology and Climatology*, vol. 54, no. 4, pp. 825–846, 2014.

[14] W. Shuang, J. Shaohui, and L. Xiaofeng, "Typhoon eye extraction with an automatic sar image segmentation method," *International Journal of Remote Sensing*, vol. 35, pp. 3978–3993, 2014.

[15] C.-J. Zhang, Q. Luo, Y. Chen, J. Lu, L.-C. Xue, and X.-Q. Lu, *Tropical Cyclone Center Determination Algorithm by Texture and Gradient of Infrared Satellite Image*, IntechOpen, Croatia, Rijeka, 2019.

[16] G. Zheng, J. Liu, J. Yang, and X. Li, "Automatically locate tropical cyclone centers using top cloud motion data derived from geostationary satellite images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 99, pp. 10175–10190, 2019.

[17] N. Jaiswal and C. M. Kishtawal, "Automatic determination of center of tropical cyclone in satellite-generated ir images," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 460–463, 2011.

[18] A. J. Wimmers and C. S. Velden, "Objectively determining the rotational center of tropical cyclones in passive microwave satellite imagery," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 9, pp. 129–136, 2010.

[19] J. Liu, C. Liu, B. Wang, and D. Qin, "A novel algorithm for detecting center of tropical cyclone in satellite infrared images," in *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, 2015.

[20] Z. Changjiang, X. Licheng, M. Leiming, and L. Xiaoqin, "Infrared brightness-temperature variance method for the objective location of tropical cyclones," *Journal of Image and Graphics*, vol. 23, no. 3, pp. 0450–0457, 2018.

[21] C. Zhao, J. Song, H. Leng, and J. Zhao, "Objective center-finding algorithm for tropical cyclones in numerical models," *Atmosphere*, vol. 10, no. 7, p. 376, 2019.

[22] X. Xu, C. Liu, and J. Liu, "A novel algorithm for the objective detection of tropical cyclone centres using infrared satellite images," *Remote Sensing Letters*, vol. 7, no. 6, pp. 541–550, 2016.

[23] J.-B. Jung, K. A. Park, K.-A. Park, D.-S. Byun, K.-Y. Jeong, and E. Lee, "Estimation of typhoon center using satellite sar imagery," *Journal of the Korean Earth Science Society*, vol. 40, no. 5, pp. 502–517, 2019.

[24] K. L. Isabella, S. Ali, L. Xiaofeng, C. T. John, and L. Zeyu, "Extracting hurricane eye morphology from spaceborne sar images using morphological analysis," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 117, pp. 115–125, 2016.

[25] S. Jin, S. Wang, X. Li, L. Jiao, J. A. Zhang, and D. Shen, "A salient region detection and pattern matching-based algorithm for center detection of a partially covered tropical cyclone in a sar image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 280–291, 2017.

[26] C. Diya, A. P. Tanvi, J. Shubham, and P. Ghanshyam, "Image segmentation using morphological operations," *International Journal of Computer Application*, vol. 117, no. 18, pp. 16–19, 2015.

[27] M. Priyanka and B. C. Bidyut, "A survey of hough transform," *Pattern Recognition*, vol. 48, no. 3, pp. 993–1010, 2015.

[28] Z. Xu and C. Qian, "A novel method for circular objects extraction based on region information," *Signal Image and Video Processing*, vol. 6, 2021.

[29] Z. Guopu, Z. Shuqun, Z. Qingshuang, and W. Changhong, "Boundary-based image segmentation using binary level set method," *Optical Engineering*, vol. 46, no. 5, Article ID 050501, 2007.

[30] H. Q. Sun and Y. J. Luo, "Adaptive watershed segmentation of binary particle image," *Journal of Microscopy*, vol. 233, no. 2, pp. 326–330, 2009.

[31] X. Zheng and T. Chen, "High spatial resolution remote sensing image segmentation based on the multiclassification model and the binary classification model," *Neural Computing & Applications*, pp. 1–8, 2021.

[32] S. Pare, A. Kumar, G. K. Singh, and V. Bajaj, "Image segmentation using multilevel thresholding: A research review," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, no. 1, pp. 1–29, 2020.

Research Article

Virtual Reality Technology Based on Embedded Image System Simulates Urban Disasters

Yuanyuan Wang and Surng Gahb Jahng

Graduate School of Advanced Imaging Science, Multimedia & Film in Chung-Ang University, Seoul, 156-756, Republic of Korea

Correspondence should be addressed to Surng Gahb Jahng; 171843186@masu.edu.cn

At present, urban disasters are generally valued by various companies and organizations. Although many companies claim to have effectively tested their urban disaster recovery plans, the effectiveness, completeness, and efficiency of their test plans still exist in many aspects. Therefore, it is very important to effectively check the city's disaster recovery plan. Nowadays, both the scientific research community and the industrial community have paid enough attention and research to urban disaster backup. This is a shield of a guaranteed simulation system. The research in this paper belongs to the category of urban disaster virtual reality technology. Urban disaster virtualization is an integral part of virtual reality technology in the current information technology field. Based on the embedded image system, this article divides the virtualization of urban disasters into two meanings: one is to virtualize various disasters into one type of disaster, so as to focus attention on the process and mechanism of disasters on the information system. The meaning is to virtualize different disaster response methods in the standard process, and all links are interrelated and interdependent so that the highest level of disaster response can be implemented. The main goal of disaster virtualization is to integrate different disasters into a comprehensive model. This paper studies urban disasters through embedded image systems and uses virtual reality technology to establish a simulation system to complete the test.

1. Introduction

At present, society pays great attention to all aspects of urban disasters, and at the same time, they have carried out pre-plans [1]. Many companies or organizations have conducted comprehensive tests on urban disaster recovery plans and have a complete set of procedures, but almost half of the tests are unsuccessful [2]. This phenomenon has led to many negative effects such as increased risks, increased costs, and impacts on corporate brands, customer experience, and customer loyalty [3]. Although many companies consider it very important to test urban disaster plans, there are still many IT personnel who have not worked out effective methods to improve the test plan, and one of the ideas for testing urban disaster recovery plans is the urban simulation system and the disaster-to-information system [4]. Since we only focus on the impact of the disaster on the information

system, rather than the disaster itself, as long as we know what impact the disaster has on the information simulation system, we can simulate the disaster [5]. In this way, the verification of the disaster recovery plan comes down to the simulation of the disaster [6]. Therefore, the urban disaster simulator designed in this paper relies on virtual reality technology to simulate the impact of fire, earthquake, and network information on the urban system [7]. It is more detailed, more comprehensive, and more effective than direct conversion to verify the disaster backup mechanism [8]. The disaster simulator is divided into three main modules: LAN IP address, disaster evolution module, and disaster on information system impact module. The LAN IP address module uses tools to find IP addresses and select nodes for simulation; the disaster evolution module uses FLEX technology to simulate disaster evolution process, dynamic simulation of the impact of time and space on the outbreak

of disasters; the disaster impact module on the information system uses the schtasks statement in the DOS statement, which is intended to connect to the target node remotely, then copy the simulated program, and finally execute it on the target node of the system the disaster simulation program is used to simulate the consequences of a disaster, so as to simulate the impact of the disaster on the information system. The main research work of this paper is to simulate the occurrence, development, and impact of urban disasters on the information system. Then, based on the embedded image system, the evacuation problem is studied [9]. This is an important safety factor in modern society and has attracted attention from all walks of life [10]. Aiming at the evacuation of people in urban disaster emergencies, combined with the uncertain factors in the evacuation process, a multiagent-based evacuation simulation model is established in a multidisaster environment in order to study the laws of people's behavior in disasters [11].

2. Related Work

The literature shows the general requirements of disaster simulators and the knowledge and technologies involved in this article, namely, JavaScript, JBoss, JSP, FLEX. The literature proposes a crowd evacuation simulation system in a multihazard environment and conducts a simulation test for the evacuation of the system [12]. The simulation results show that the system better demonstrates the impact of the disaster environment on the evacuation of the people, and the simulation effect on the evacuation is more realistic [13]. It embodies the scientificity and rationality of the crowd evacuation model and establishes a mathematical model based on current disasters [14]. Fire, explosion, and the spread of irritating toxic smoke are the main disasters. This model focuses on the impact of fire and explosion sound on human health and establishes a human perception model and makes it as true as possible based on the actual location of each person. The information is imitated that people can perceive in the environment. The literature shows that FEKO creates a true and false target model in the SAR image, and according to the simulation results, starting from the scattering mechanism, it analyzes the effect of the difference in the geometric structure of the true and false target on the echo scattering [15]. According to the analysis result, a method is proposed. Combining the target feature extraction algorithm of gray and texture, the validity of SAR image authenticity target recognition is determined. The literature shows the definition and characteristics of virtual reality technology and its application advantages in urban disasters, as well as the definition of disaster process exploration mode and disaster processing mode steps [16]. A lot of research on the theoretical basis of virtual city disaster platform research has also been carried out. The literature shows the concept and meaning of unit testing, integration testing, and system testing, and lists some test cases, evaluates the final test results, and compares them with similar solutions and similar products, and finally shows the development and operating environment of the system.

3. Embedded Image System and Virtual Reality Technology

3.1. Embedded Image System

3.1.1. Embedded System. The development of the embedded system from centralized to a system application platform, open system structure, resource service, intelligent management decision, and green direction is not only a technical upgrade but also meets the needs of practical applications. The system architecture is not only highly modular but also supports reconstruction, modification, and expansion. Chinese universities and scientific research institutes have accumulated some research foundations in embedded systems and reliability calculations, and some scientific research institutions have accumulated rich research experience in embedded systems and microprocessors. Based on the strategy of vigorously developing the advanced manufacturing industry, China has carried out embedded system research in the industrial field and puts forward the requirement of combining intelligence and industrialization. The equipment manufacturing industry is the basic requirement for industrialization. Hardware and software are used in the design of CNC machine tools in the software industry. It is very important in the manufacture of machine parent equipment. CNC system is a typical application of the embedded real-time system in the industrial field.

In foreign countries, the design of embedded systems is also a matter of close concern in academia and industry. Some countries have incorporated the research of embedded systems into the national key strategic research plans and support related research centers.

Now, in the international manufacturing industry, mechanical or electrical components are completely embedded in controlled equipment to create a dedicated computer system designed for specific applications. In the future, embedded systems will develop into cyber-physical systems (CPS) in the industrial field. There will be a large amount of data exchange between sensors, actuators, and computing devices, with the purpose of adapting to different application environments. Data interaction and cooperation are affected by the uncertain factors of the physical world, and it is also a new challenge to the reliability of the system. The design of embedded systems is more inclined to the requirements of low power consumption and higher reliability.

3.1.2. Image Simulation. In the image simulation process, when electromagnetic waves are reflected on the target, a random elevation difference is added to the diameter of the tracking tube to simulate the scattering characteristics of the target surface. On this basis, the combined electromagnetic scattering mode of the target in a complex scene is established as follows:

$$P'_j = P_j + \Delta H_j \cdot \widehat{n}_i. \tag{1}$$

On the surface of the rail pipe with random elevation, the normal direction of the original reflecting surface is

corrected, and $\hat{n}\_\hat{i}$ is the normal direction of the triangle facet after the elevation difference is introduced.

$$\hat{n}'_i = \frac{(P'_2 - P'_1) \times (P'_3 - P'_1)}{(P_2 \cdot -P'_1) \times (P_3 \cdot -P'_1)}. \tag{2}$$

Rayleigh distribution:

$$P_r(x) = \frac{2x}{\alpha} \exp\left(-\frac{x^2}{\alpha}\right). \tag{3}$$

Lognormal distribution:

$$P_l(x) = \frac{1}{x(\pi\alpha)^{1/2}} \exp\left(-\frac{(\ln x - m)^2}{\alpha}\right), \tag{4}$$

where m and α represent the mean and variance of lnx, respectively, which can be obtained by maximum likelihood estimation:

$$\left\{ \hat{m} = \frac{1}{N} \sum_{i=1}^{N} \ln(X_i)\hat{\alpha} = \frac{2}{N} \sum_{i=1}^{N} [\ln(X_i) - m_{ML}]^2. \tag{5}\right.$$

Weibull distribution:

$$P_w(x) = \frac{\gamma}{\omega}\left(\frac{x}{\omega}\right)^{\gamma-1} \exp\left[-\left(\frac{x}{\sigma}\right)^{\gamma}\right] \tag{6}$$

It can be obtained by the maximum likelihood estimation method:

$$\left\{ \begin{array}{l} \hat{\gamma} = \left\{ \frac{6}{\pi^2} \frac{N}{N-1} \left[\frac{1}{n} \sum_{i=1}^{N} (\ln(X_i))^2 - \left(\frac{1}{n} \sum_{i=1}^{N} \ln(X_i) \right)^2 \right] \right\}^{1/2} \\ \\ \hat{\omega} = \exp\left[\frac{1}{N} \sum_{i=1}^{N} \ln(X_i) + 0.5772\hat{\gamma}^{-1} \right] \end{array} \right. \tag{7}$$

K distribution:

$$P_k(x) = \frac{2}{\alpha\Gamma(\mu+1)}\left(\frac{x}{2\alpha}\right)^{\mu+1} K_\mu\left(\frac{x}{\alpha}\right). \tag{8}$$

In the case of a large number of sea clutter samples, the second-order and fourth-order methods are the most effective. The results are as follows:

$$\left\{ \begin{array}{l} \hat{\mu} = \frac{(25/16 - M_{5/2}/M_{1/2}M_2)}{(M_{5/2}/M_{1/2}M_2 - 5/4)} \\ \\ \hat{\alpha} = \frac{M_1}{\sqrt{\pi}\,\Gamma(\mu+1)\Gamma(\mu+1.5)} \end{array} \right., \tag{9}$$

where

$$M_n = \frac{1}{N} \sum_{i=1}^{N} X_i^n. \tag{10}$$

is the central moment of order n of the statistic.

The KS test method is a test method to test whether the data distribution is consistent with the commonly used theoretical distribution. The KS value is the maximum value of the difference between the cumulative empirical data distribution function and the expected cumulative distribution function, denoted by De, used to test the statistics of sample data distribution, where De is expressed as

$$De = \max_{1<i<l}|F(x_i) - F_l(x_i)| = \max_{1<i<l}\left|\frac{n}{N} - F_l(x_i)\right|. \tag{11}$$

KL distance is a measure of information theory. If the assumed distribution is $F(x_i)$ and the actual distribution is $F_l(x_i)$, the expression of the measure is

$$D(F_l F) = \int_{i=1}^{N} F_l(x_i)\log_2\left(\frac{F_l(x_i)}{F(x_i)}\right)dx_i. \tag{12}$$

For the discrete case, the expression can be simplified to

$$D(F_l F) = \sum_{i=1}^{N} F_l(x_i)\log_2\left(\frac{F_l(x_i)}{F(x_i)}\right). \tag{13}$$

The detection mean square error can be defined as

$$mse = \sum_{i=1}^{N} \frac{[F(x_i) - F_l(x_i)]^2}{N}. \tag{14}$$

Sum the data of all subsegments and average them to obtain the power spectrum of the signal to be estimated, which can be expressed as

$$\overline{P}_{per}(\omega) = \frac{1}{L} \sum_{i=1}^{L} \hat{P}^i_{per}(\omega). \tag{15}$$

The removal of phase imbalance can be removed by the following formula:

$$I_{rot} = [I - Q^* \sin(\beta)]sqrt(1 - \sin(\beta)^2). \tag{16}$$

The actual parameters of the measured scene are sea temperature 7.6°C, wind speed 5.3 m/s, 0° is true north, wind direction 300°, radar appearance direction 125.6°, maximum sea wave height at sea 0.8 m, distance direction 2575–2768 meters, a total of 196 meters, the distance resolution is 30 m, and the target distance tar A is in the seventh unit. The seventh distance unit data is selected as the analysis object. Figure 1 shows the amplitude-time characteristics of the seventh unit under the same data polarity. The maximum amplitude under VV polarity is 39.38, and the maximum amplitude under HH polarity is 111.5.

Figure 2 shows the statistical probability density function of the echo amplitude.

Tables 1 and 2 respectively, check the distribution of the distance unit clutter under VV polarity and HH polarity. From the mean value test, under the condition of VV polarization, the clutter amplitude is very suitable for the K distribution. In the case of transformation, the amplitude of the clutter fits well with the Weibull distribution.

The clutter distribution test of the seventh distance unit of Document 54 is shown in Table 2:

3.2. Virtual Reality Technology. When a fire occurs, the nearby temperature rises and produces flames, smoke particles, and toxic gases. The harmful gases produced by the
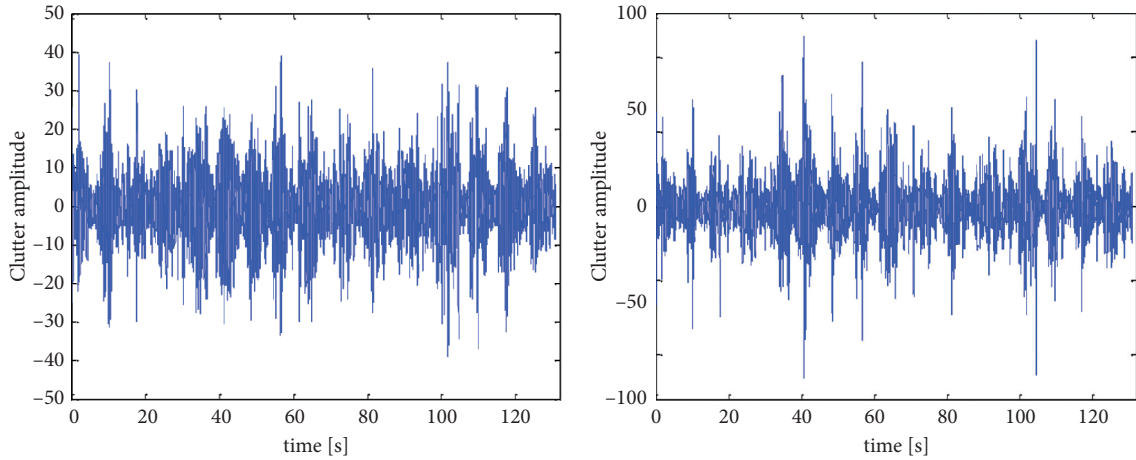
FIGURE 1: The time series of the amplitude of the seventh range unit with the same polarization.
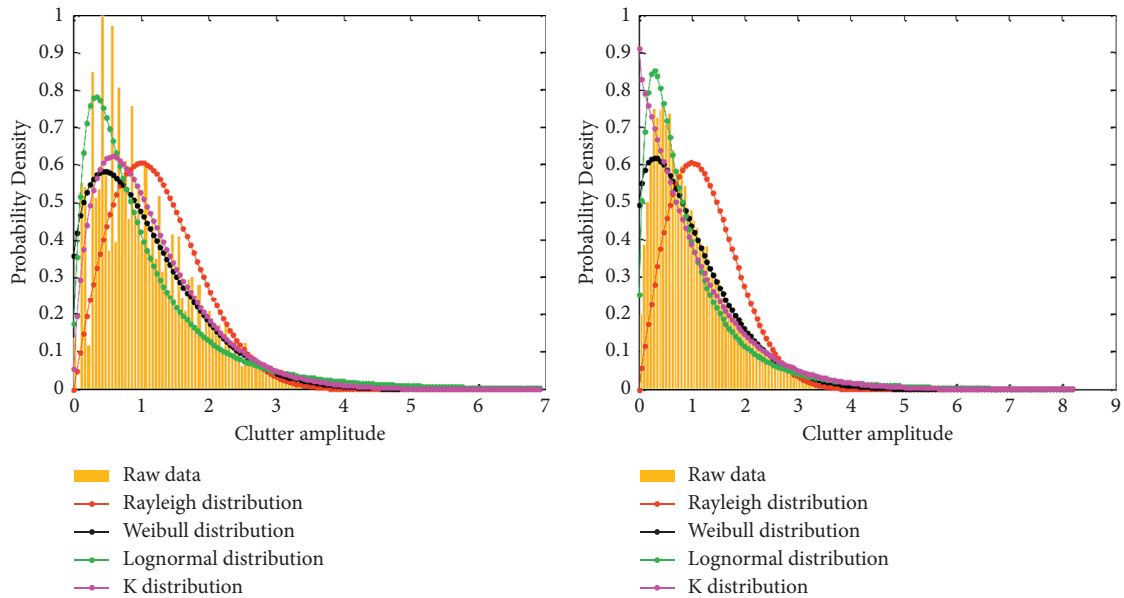
FIGURE 2: The distribution fitting of the amplitude of the seventh distance unit with the same polarization.

TABLE 1: Clutter distribution test of the seventh distance unit (VV polarization).

---	Rayleigh distribution	Weibull distribution	Lognormal distribution	K Distribution
K-S inspection	0.9247	0.9247	1.7646	0.7875
K-L inspection	13.0163	17.1023	17.1412	8.9692
MES inspection	0.0171	0.0083	0.0097	0.0073
Test mean	3.4896	4.5089	4.7289	2.4408

TABLE 2: Clutter distribution test of the seventh distance unit (HH polarization).

---	Rayleigh distribution	Weibull distribution	Lognormal distribution	K Distribution
K-S inspection	0.5978	0.5889	0.9922	1.5817
K-L inspection	14.7828	0.8135	0.8952	2.0242
MES inspection	0.0162	0.0019	0.0015	0.0067
Test mean	3.8493	0.3534	0.4722	0.9032

fire have psychological and physiological effects on the evacuation of people. Since there are many substances produced by fire in practice, the impact on evacuation may be more complicated. Only by simplifying the factors considered in the fire can a corresponding fire model be established. The size of the fire is determined by the power of

the fire source, which shows the area of the flame. The diameter D and height L of the bottom of the fire range can be calculated by

$$D = \sqrt{\frac{4Q}{\pi Q_e}},$$

$$H = -1.02\,D + 0.235Q^{2/5}. \tag{17}$$

After a fire occurs, smoke is often produced and diffused to the edge, and its diffusion rate is generally determined according to the gas molecule velocity distribution function.

$$V = 1.6\sqrt{\frac{RT}{M}}. \tag{18}$$

At the initial moment, the concentration and temperature distribution in the air mass are uniform, and the gas concentration is calculated according to the stable continuous point source smoke flow model.

$$C(x, y, h, H) = \frac{Q}{2\pi u \sigma_y \sigma_z} \exp\left(-\frac{y}{2\sigma_y^2}\right) \cdot \left\{\exp\left[\frac{-(h-H)^2}{2\sigma_z^2}\right] + \exp\left[\frac{-(h+H)^2}{2\sigma_z^2}\right]\right\}. \tag{19}$$

The steady-state sound pressure level formula is used to calculate the sound pressure level at a certain position.

$$L_p = L_w + 10\lg\left(\frac{Q}{4\pi r^2} + \frac{4}{R}\right). \tag{20}$$

During the evacuation process, fires often produce smoke and other toxic gases, and other irritating odors may also appear. Pedestrians can perceive the existence of such irritating gases through their sense of smell during travel, which affects people's evacuation routes to a certain extent. Judgment, it also causes certain harm to pedestrians. After a gas leak occurs, people can use the location information and the above equation (18) to calculate the corresponding gas concentration. By obtaining the gas concentration within the range of the person, the human sense of smell can be realized. If the leaked gas is toxic, it will have a significant impact on human physiology. Assuming that the dose of poisonous gas a person bears is S and the probability of causing harm to the person is Ps, the probability formula can be calculated as follows:

$$P_s = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\gamma-5} \exp\left(-\frac{u^2}{2}\right) du. \tag{21}$$

Among them, the random variable Y can appear in the form of a nonlinear equation. When the toxin concentration does not change with time, its toxin value can be calculated as follows:

$$Y = k_1 + k_2 In(C^n t). \tag{22}$$

From this, a simple calculation formula for the degree of physiological influence β can be calculated as follows:

$$\beta = \frac{P_s + \theta}{\theta}. \tag{23}$$

This article derives a rough calculation following formula for the degree of psychological influence based on empirical data:

$$\alpha = \left(\frac{L_{left} + L_{right}}{2} - 120\right) \cdot \sigma. \tag{24}$$

4. Urban Disaster Simulation System Design and Application Analysis

4.1. System Characteristics and Target Analysis

4.1.1. System Analysis. The disaster simulator is mainly used to remotely control the local information system, simulate the impact of the disaster on the information system, and simulate the consequences of the disaster. The key question is how to simulate the impact of the disaster on the information system. In order to accurately simulate the disaster effect, this paper divides the disaster simulator into three modules in the design process: scan to obtain IP address, disaster development module, and simulate the impact of disaster on nodes. First, the network is scanned to obtain the IP address module to achieve the simulation target and delimit the simulation range; that is, the simulator starts; then, the disaster development module can display the changes of the disaster over time and space, truly restore the occurrence and evolution of the disaster, and finally simulate the disaster. The node impact module is the core of the disaster simulator. The simulator creates three states of the information system so that it can truly simulate the impact of the disaster on the information system. The three states are error, malignant change, and crash. Errors indicate the system's failure. The calculation result is incorrect. Malignant change means that the system is running in a random state; crash means that the system is down.

The simulator simulates the impact of disasters on the information system, but only as a simulation. Therefore, security must be considered, and no irreversible impact on the system must be taken into account. In this way, efficiency and security can be achieved at the same time, and a real simulation effect can be achieved.

In addition to its functions, the system must also meet the following characteristics:

Security: Because it is only a disaster simulation, it will not have an irreversible impact on the system. For programs that attack the system.

Clean up after the simulation.

Convenience: Easy to implement and run, the simulator is flexible and convenient, and user operation is simple.

Scalability: Due to technological progress and continuous changes in business conditions, this simulator must have a certain degree of scalability, which can be further expanded and improved.

Efficiency and safety are the basic requirements of this disaster simulator. Convenience and scalability are the

highest requirements for the disaster simulator and are the guarantee for improving the quality of the disaster simulator.

After analyzing the functions listed in the requirements, the simulator can be divided into 3 modules.

Automatically, the network is scanned to find the IP address: to obtain the IP address in the local network, you can select the node that simulates the disaster and enters its physical location. Here, you must select the node that needs to be simulated, or you can select all, because sometimes users do not want all the nodes on the local area network. The machines are simulated, so you can choose.

Disaster propagation time and space effects are simulated: the type, intensity, and environmental factors of the disaster are selected to simulate the disaster propagation process. When the disaster spreads to a certain node, the simulator will simulate the impact of the disaster on the system on that node.

Simulation of the impact of disasters on nodes: The consequences of disasters on simulated nodes are realized through remote control. This module is the basic part of the disaster simulator, which can simulate the occurrence, development, and impact of disasters on information systems in detail and truly.

4.1.2. The Goal of the System. The basic use of the disaster simulator is to detect the disaster backup mechanism by simulating the impact of fire, earthquake, and disaster virus on the information system and to simulate the occurrence of disaster and the impact on the information system.

In order to achieve the goal of the disaster simulator to simulate the impact of the disaster on the information system, the simulator must achieve the following effects:

(1) Detailed and realistic simulation of the occurrence, development, and impact of disasters on information systems

(2) Check the status of the information system after the attack, and check the stability of the system

(3) Check the backup mechanism of the information system and provide data for user analysis

(4) Check the database backup mechanism, which can be analyzed and improved by the database administrator

(5) Check the stability and security of the network

(6) Test all users' disaster recovery plans. Only when these overall effects are achieved, the goal of the disaster simulator can be achieved

Many staff members participated in the operation, operation and cooperation of the disaster simulator, the main relevant participants: users, user system administrators, network administrators, database administrators, user system administrators, and disaster backup personnel. The following is a brief introduction of the main personnel.

User: The user needs to use the disaster simulator correctly according to the operation manual, input various effective parameters of the simulation, and clean up the simulated program according to the operation manual.

System maintenance personnel: System maintenance personnel not only analyze its related data but also check the impact on the system.

Network administrator: The network administrator is mainly responsible for two parts: checking and setting up the network, ensuring the normal use of the emulator, and checking the security and stability of the network according to the attack effect of the emulator.

Database administrator: Pay attention to the impact of the simulator on the database. The stability and backup mechanism of the database should be carefully checked.

User system administrator: The user system administrator needs to analyze and check the effect of the simulator, especially to check its security and stability.

Disaster backup personnel: When analyzing the entire operation process, disaster backup personnel monitors and improves their disaster backup mechanism.

The development and design of disaster simulators need to fully consider relevant personnel so that the effects of disaster simulators are more perfect.

4.2. Overall System Design. As a traditional tool, the flow chart of the disaster system is mainly used to describe the physical model of the system. It is a graph that uses some graphic symbols to describe the process of the system. It is required to describe various components, such as tables, files, and programs, and display them.

The function of the flow chart of the disaster system is manifested in the following aspects: The flow chart of the disaster system is a tool for business operators, administrators, and system analysts to communicate with each other. Through the flow chart of the entire disaster system, members can communicate well. Another function of the system flow chart is to allow the system analyst to effectively understand the overview of the disaster system business processing, which is also the basis for the disaster system analyst to make further analysis, and it is also the beginning of the system design.

The flowchart system not only understands the entire disaster system but can also analyze the rationality of the business process of the disaster system.

From the analysis of the requirements of disaster software to the completion of the final development, the most important step is the use case diagram. The use case diagram describes how people use the disaster system and shows the relationship between users and what services users need so that users of the system can understand more easily. The purpose of these elements is also convenient for software developers to finally realize these elements. Use case diagrams are generally used by various development activities, many of which are used to describe systems and subsystems. The system can be divided into very specific units, so it has become an excellent tool for software developers.

As shown in Figure 3, this system mainly has six use cases.

Figure 4 shows the interaction events between the users of the system and the disaster system: first, the system is started, and then, the system returns all the IP addresses of the local area network, and then, the user selects the node,
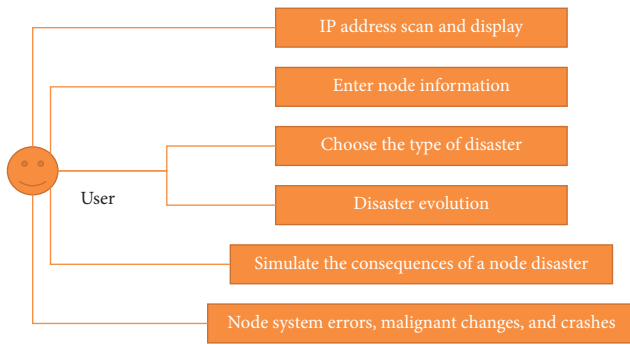
FIGURE 3: System use case diagram.
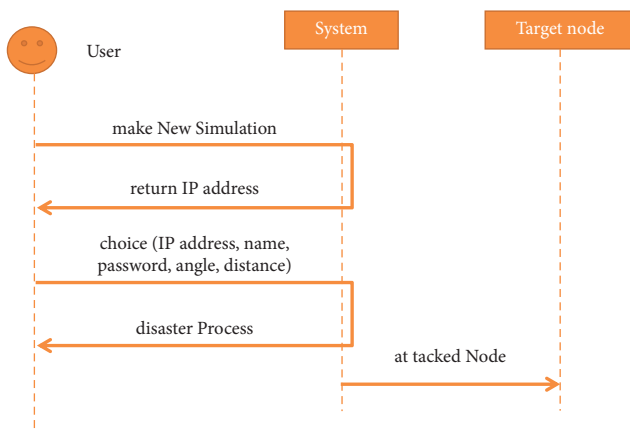
FIGURE 4: System sequence diagram.

the disaster type, and environmental factors to simulate, and finally, the system simulates the evolution process of the disaster. The system simulates the consequences of implementing a disaster on the node.

The system sequence diagram is a fast and simple product that illustrates the input and output events related to the system in question. The system sequence diagram is the input to the operation contract and object design, which shows the external participants directly interacting with the disaster system, and the participants initiated system events. For some use cases, system sequence diagrams should be drawn for the main success scenarios of each use case, as well as frequently occurring or complex alternative scenarios.

4.3. System Function Design. Receiving IP address module: There are four methods for receiving IP addresses, DNS, SNMP, ping command, and Traceroute command, but these contents are not the key issues of the disaster simulator, and we can obtain IP addresses by using tools. There are many such tools and found that these tools are more complicated and confusing, not suitable for disaster simulators. Finally, this article found a simple and practical tool. Local area network IP auto-configurer, its interface, and operation are relatively simple and can be used in the disaster simulator designed in this paper. This module is to start the disaster simulator, find all the IP addresses on the device LAN, and then select the node to be simulated to achieve complete

simulation, and the distribution map of the target node can be formulated according to the node selected by the user and input parameters for later dynamic demonstration.

Disaster evolution module: First, you need to use animation technology to simulate the evolution of fires, earthquakes, and computer viruses and make models. The simulator uses FLEX to achieve the simulation effect. Flex uses a non-Flash path to interpret the .mxml files to organize components and create corresponding files. Among them, this article uses FLEX technology to simulate the time and space effects of the disaster spread on the page, that is, dynamically simulate the evolution process of disasters. After the user selects the parameters, the background program formulates the evolution formula according to the parameters, and the page changes according to the evolution simulation to simulate the change of the disaster. For fires, earthquakes, and computer viruses, different evolution models are created to adapt to the different nature of disasters and try to truly simulate the evolution process of disasters in the local area network in reality.

System impact module: The core of the disaster simulator is the system impact module. To start using this module, you need to complete the following steps: (1) To set up a link to each target node, you must first be able to connect to each target node; (2) to create an empty folder in the designated area of the target node, it is best to create this folder on the C drive, because most computers have a C drive; (3) to copy the simulation program to an already created folder; (4) to remotely start the simulation program. The simulation program is mainly composed of bat files, VBScript programs, and enhanced computer code viruses. The process of system error, system malignancy, and system crash must be realized. These steps must be controlled by the background program, adapt to the disaster evolution process, and be able to cover all target nodes. This module is essential for simulating disasters. It must be able to accurately and timely implement the simulation, and it must realistically simulate the consequences of the disaster. Therefore, the simulation program must be designed to be effective and powerful.

The disaster evolution structure is shown in Figure 5.

4.4. System Test Analysis. In the development process of the entire disaster system, unit testing is the lowest level of testing. When testing a unit, the unit that needs to be tested is isolated from other parts. Unit testing is repeatable. Test units can be tested repeatedly to make the test results more accurate. Therefore, the maintenance of unit testing cannot exceed the life cycle of all software.

The unit test case style is shown in Table 3.

The main purpose of integration testing is to combine different modules to test to find problems and defects in each interface. If data are not transmitted through the interface, the interface does not have a connection effect; the two modules are integrated together and affect each other, resulting in failure of function realization; after the integration, the module did not achieve the expected function and so on.
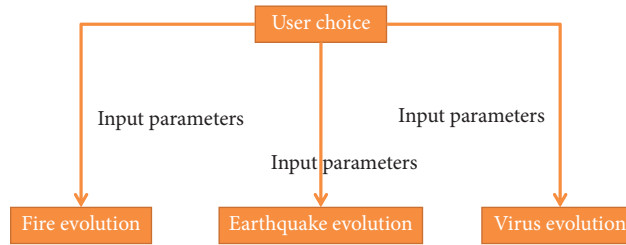
FIGURE 5: Disaster evolution structure diagram.

TABLE 3: Test case style.

Test case number	Test-01
Test item	Disaster simulator
Test title	Test of the disaster simulator
Importance level	High
Parameter input	Test
Execution steps	—
Expected result	Achieve the desired goal

TABLE 4: Test case IPA-001.

Test case number	diSimulator-IPA-001
Test item	Disaster simulator
Test title	Perform integration test on the module for obtaining the IP address
Importance level	High
Prefabricated conditions	—
Parameter input	Start IP address, end IP address
Execution steps	Automated tool execution
Expected result	Obtain all IP addresses in the selected segment of the LAN

TABLE 5: Test case T-001.

Test case number	diSimulator-T-001
Test item	Disaster simulator
Test title	System test of disaster simulator
Importance level	High
Prefabricated conditions	Select multiple simulation nodes
Parameter input	Choose fire, earthquake, computer virus separately, choose different intensity, choose different environmental factors
Execution steps	—
Expected result	All selected nodes eventually crash

The IP address module integration test case is obtained, as shown in Table 4.

System testing is a combination of software, hardware, external equipment, network, interface, and other elements required for system operation, and then, the overall test is performed to verify whether the system functions, performance, and functions meet the requirements of the entire system and to understand which areas are not. The requirements are met, and targeted changes and improvements are made. After testing the system and discovering the problem, it is necessary to be able to locate the problem and know where the problem is, so as to facilitate the improvement of the problem. The system test requirements can cover all the main functions, including not only the test

software but also other components such as hardware, external equipment, and networks, interfaces.

Disaster simulator system test cases are shown in Table 5.

5. Conclusion

This paper focuses on the simulation of evacuation in an urban disaster environment based on virtual reality technology and establishes an evacuation awareness model. Compared with the research on evacuation based on perception patterns at home and abroad, this paper combines people and the environment and combines the environment and perception evacuation research from the perspective of human perception, and expounds the impact of disasters on

evacuation. It also simulates the impact of fires, earthquakes, and information systems on the city. According to the analysis of the impact of urban disasters on the urban information system, the entire framework is designed. The simulator is divided into three functional modules: IP address acquisition, disaster evolution, and the impact of the system. The three modules were designed, developed, and tested in sequence, the three modules were integrated, and the system was tested, and a disaster simulator was deployed. In order to better respond to urban disasters, it is necessary to improve the recommendations and norms of disaster response, and it is urgent to conduct disaster simulation research. This paper has done a lot of research on fire, earthquake, and information system simulation, and this direction will inevitably become a new and valuable research direction.

Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, after the publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Pan, C. S. Han, K. Dauber, and K. H. Law, "A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations," *Ai & Society*, vol. 22, no. 2, pp. 113–132, 2007.

[2] M. Mahootchi and S. Golmohammadi, "Developing a new stochastic model considering bi-directional relations in a natural disaster: a possible earthquake in Tehran (the Capital of Islamic Republic of Iran)," *Annals of Operations Research*, vol. 269, no. 1-2, pp. 439–473, 2017.

[3] S. H. Kang and M. Skidmore, "The effects of natural disasters on social trust: evidence from South Korea," *Sustainability*, vol. 10, no. 9, pp. 1–16, 2018.

[4] N. Hooshangi and A. Asghar Alesheikh, "Agent-based task allocation under uncertainties in disaster environments: an approach to interval uncertainty," *International Journal of Disaster Risk Reduction*, vol. 24, pp. 160–171, 2017.

[5] V. Grimm, E. Revilla, U. Berger et al., "Pattern-oriented modeling of agent-based complex systems: lessons from ecology," *Science*, vol. 310, no. 5750, pp. 987–991, 2005.

[6] M. Hashemipour, S. Stuban, and J. Dever, "A disaster multiagent coordination simulation system to evaluate the design of a first-response team," *Systems Engineering*, vol. 21, no. 4, pp. 322–344, 2018.

[7] M. Janssen, J. Lee, N. Bharosa, and A. Cresswell, "Advances in multi-agency disaster management: key elements in disaster research," *Information Systems Frontiers*, vol. 12, no. 1, pp. 1–7, 2010.

[8] A. Aydeger, N. Saputro, K. Akkaya, and S. Uluagac, "SDN-enabled recovery for Smart Grid teleprotection applications in post-disaster scenarios," *Journal of Network and Computer Applications*, vol. 138, pp. 39–50, 2019.

[9] S. Iwanaga and A. Namatame, "Contagion of evacuation decision making on real map," *Mobile Networks and Applications*, vol. 21, no. 1, pp. 206–214, 2016.

[10] B. Su, H. Huang, and Y. Li, "Integrated simulation method for waterlogging and traffic congestion under urban rainstorms," *Natural Hazards*, vol. 81, no. 1, pp. 23–40, 2016.

[11] Q. Yang, Y. Sun, X. Liu, and J. Wang, "MAS-based evacuation simulation of an urban community during an urban rainstorm disaster in China," *Sustainability*, vol. 12, no. 2, pp. 1–19, 2020.

[12] M. Nasir, S. Nahavandi, and D. Creighton, "Fuzzy simulation of pedestrian walking path considering local environmental stimuli," in *Proceedings of the 2012 IEEE International Conference on Fuzzy Systems*, pp. 1–6, IEEE, Brisbane, Australia, June 2012.

[13] Y. Han and H. Liu, "Modified social force model based on information transmission toward crowd evacuation simulation," *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 499–509, 2017.

[14] C. Feliciani and K. Nishinari, "An improved Cellular Automata model to simulate the behavior of high density crowd and validation by experimental data," *Physica A: Statistical Mechanics and its Applications*, vol. 451, pp. 135–148, 2016.

[15] O. Yousif and Y. Ban, "Improving urban change detection from multitemporal SAR images using PCA-NLM," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 2032–2041, 2013.

[16] Z. Lv, D. Chen, R. Lou, and H. Song, "Industrial security solution for virtual reality," *IEEE Internet of Things Journal*, vol. 99, p. 1, 2020.

Retraction

Retracted: A Study of Feature Construction Based on Least Squares and RBF Neural Networks in Sports Training Behaviour Prediction

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] C. Qiu, C. Su, X. Liu, and D. Yu, "A Study of Feature Construction Based on Least Squares and RBF Neural Networks in Sports Training Behaviour Prediction," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5034081, 7 pages, 2022.

Research Article

A Study of Feature Construction Based on Least Squares and RBF Neural Networks in Sports Training Behaviour Prediction

Chunyan Qiu,[1] Changhong Su,[2] Xiaoxiao Liu,[3] and Dian Yu [1]

[1]*College of Sports and Physical Education, Shandong Sport University, Rizhao 276827, Shandong, China*
[2]*Sports Section, Guangshui Road Primary School, Qingdao 266100, Shandong, China*
[3]*Faculty of Engineering and Technology, School of Engineering, Liverpool John Moores University, Liverpool, UK*

Correspondence should be addressed to Dian Yu; yudian@sdpei.edu.cn

This paper examines the problem of athletes' training in sports, exploring the methods and means by which athletes can perform difficult movements in which they normally make minor training errors in order to achieve better competition results and placements. To this end, we test the explanatory and predictive effects of a theoretical model starting with planned behaviour and then use exercise planning, self-efficacy, and support as variables to develop a partial least squares regression model of sports to improve the explanation and prediction of sporting athletes' intentions and behaviour. An improved RBF network-based method for player behaviour prediction is proposed. On the basis of the RBF analysis, the number of layers and the number of neurons in the hidden layer of the network are adjusted and optimised, respectively, to improve its generalisation and learning abilities, and the athlete behaviour prediction model is given. The results demonstrate the advantages of the improved algorithm, which in turn provides a more scientific approach to the current basketball training.

1. Introduction

What if, after countless hours of practice, a sportsperson realises that there is only a 50% chance of successfully completing one of the most difficult manoeuvres of the entire competition and that the athlete might miss the most important and critical moment of the competition? What does the athlete do when faced with these potential threats? Many scholars have suggested adjusting one's breathing to regulate one's emotions, controlling oneself from thinking that the action has a high probability of being wrong, adjusting one's mental state so that he can perform the technique successfully, etc. [1]. However, numerous studies have shown that the use of intentions alone is not a good predictor of behaviour change. Action plans and coping plans are self-adjustment strategies. In many domains, behavioural intentions are supported after planning and execution as a motivation is received. The addition of a theory of planned behaviour appears to be the most appropriate as the theory of planned behaviour lacks sequential

structure. It has also been found that when using behavioural interventions, research has not found an effect on behaviour by making implementation intentions, and on the basis of existing research, planning is more important as a mediating variable, and this paper will continue to examine and explore the forecasting effects of planning as an intermediate variable [2].

Currently, most of the research studies in the field of video-based motion human analysis are focused on motion behaviour recognition, but not on motion behaviour prediction. However, in real life, prediction of human movement based on the video is more valuable than behaviour recognition [3]. The public security department can be alerted to the predicted criminal behaviour in time to take corresponding measures. This not only reduces the amount of manpower and material and financial resources spent on security investigations but also provides effective prevention of sudden crime [4]. In the sports sector, a comprehensive set of video-based human movement prediction technologies can be used to accurately capture the game data of some

of the best teams during sports training and to create a predictive discriminator for each athlete [5, 6].

The contributions of this paper are as follows:

We start from the planned behaviour to test the interpretation and prediction effect of the theoretical model and then take the exercise plan, self-efficacy, and support as variables to establish the partial least squares regression model of sports, so as to improve the interpretation and prediction of aerobics athletes' intention and behaviour.

We propose a team member behaviour prediction method based on the improved RBF network. The RBF network is analyzed. Based on the RBF analysis, the number of network layers is adjusted and optimized to improve its generalization and learning abilities, and the athlete behaviour prediction model is given.

Based on the analysis of the sports competition video, the improved RBF is used to predict the athletes' behaviour, and the prediction results are compared with the traditional prediction results. The results show that our algorithm has advantages.

2. Related Work

At present, the research of individual sports is mainly reflected in the movement recognition and analysis of athletes [7–9] using target detection and tracking technology, the moving target is extracted from the diving video, and the moving target is video synthesized. A motion recognition method based on the human contour curvature scale space template is proposed and applied to the motion recognition of players in tennis competition [10]. The movements of ice hockey and football players are analyzed and studied. The movement direction of players is identified by using the histogram of oriented gradient (HOG) and hidden Markov model (HMM) [11–13]. The player's gestures in the low-resolution tennis game video are recognized [14]. By tracking the key points of some body parts of ballet dancers, the dance trajectory is obtained, and the automatic analysis of dance movements is realized [15]. It realizes the recognition of four swimming strokes in the swimming competition video.

In recent years, as research on individual sports analysis has matured so has research on multiplayer team sports [16]. Zhong and Xiao [17] investigated the patterns of passes that help to score goals in football and used multiscale analysis to cluster passes hierarchically and to mine the pattern categories of passes; Chen et al. [18] classified the trajectory and landing point of the ball as special cases in tennis and used an improved Bayesian network to fuse these two features. Li and Zhang [19] proposed a method for two-sided team possession situations, which is based on the ball position and

the detection of the interplay, the local minima of the ball motion trajectory velocity of a player's possession start, the local minima of the ball position coordinates, the player's position, and the ball acceleration start point to infer the player's touch point. Li et al. [20] integrated a variety of information on the pitch in the ritual to achieve automatic analysis of football game tactics.

3. Development of a Regression Model for Sports Training

Simplify the data structure and correlation analysis of bivariate variables, greatly facilitating multivariate data analysis [21]. It uniquely suits to deal with small sample sizes and independent variables with significant multiple correlations between variables. However, the partial least squares regression model can be expressed as the original regression equation for all variables. The partial least squares regression method selects principal components that also include all variables and does not fully address the more severe multivariate correlations, especially in the case of some independent variables and small sample conditions [22].

In this paper, data from sports athletes regarding physical training were used to conduct partial least squares regression modelling in terms of the explanatory and predictive power of three variables—exercise program, self-efficacy, and social support—on exercise intention and behaviour and their role in the intention-to-behaviour process. The subjects of this paper were 10 players of a sports team. The first group of variables was exercise planning, self-efficacy, and social support. In this study, social support consisted of two types, family social support and friend social support, which were not effective predictors of exercise intention. The second group of variables was the training performance indicators of the sport players, which were rotation, bending, and high jump. The raw data are shown in Table 1, and the correlation coefficient matrix is shown in Table 2.

The simple correlation matrix of these six variables shows that an exercise program and self-efficacy are positively correlated; work program and self-efficacy are negatively correlated with social support, and rotation, bending, and high jump are positively correlated.

The partial least squares regression model is calculated by presenting two components t_1 and t_2, which have a cross-validity of $Q_2^2 = -0.1969$. The regression equation for the standardised variable $\overline{y_k}$ on component t_1 in this paper is as follows:

$$\overline{y_k} = r_{1k}t_1 + r_{2k}t_2, k = 1, 2, 3. \tag{1}$$

The resulting partial least squares regression equation for component t_1 is

$$\overline{y_k} = r_{1k}\left(w_{11}^*\overline{x_1} + w_{21}^*\overline{x_2} + w_{31}^*\overline{x_3}\right) + r_{2k}\left(w_{21}^*\overline{x_1} + w_{31}^*\overline{x_2} + w_{32}^*\overline{x_3}\right)$$
$$= \left(r_{1k}w_{11}^* + r_{2k}w_{12}^*\right)\overline{x_1} + \left(r_{1k}w_{21}^* + r_{2k}w_{22}^*\right)\overline{x_2} + \left(r_{1k}w_{31}^* + r_{2k}w_{32}^*\right)\overline{x_3}. \tag{2}$$

TABLE 1: Physical training data.

No.	Exercise program	Self-efficacy	Social support	Rotate	Bend	High jump
1	91	36	50	5	162	60
2	89	37	52	2	110	60
3	93	38	58	12	101	101
4	63	35	62	12	105	37
5	89	35	46	13	155	58
6	82	36	56	4	101	42
7	67	34	60	6	125	40
8	76	31	74	15	200	40
9	54	33	56	17	251	250
10	69	34	50	17	120	38

TABLE 2: Correlation coefficient matrix.

	x_1	x_2	x_3	y_1	y_2	y_3
x_1	1	0.8702	−0.3658	−0.3897	−0.4931	−0.2263
x_2	0.8702	1	−0.3529	−0.5522	−0.6456	−0.1915
x_3	−0.3658	−0.3529	1	0.1506	0.225	0.0349
y_1	−0.3897	−0.5522	0.1506	1	0.6957	0.4958
y_2	−0.4931	−006456	0.225	0.6957	1	0.6692
y_3	−0.2263	−0.1915	0.0349	0.4958	0.6692	1

The calculations for $r_h = (r_{h1}, r_{h2}, r_{h3})$ are

$$\overline{y_1} = -0.07\overline{x_1} - 0.489\overline{x_2} - 0.122\overline{x_3},$$
$$\overline{y_2} = -0.15\overline{x_1} - 0.5244\overline{x_2} - 0.054\overline{x_3}. \tag{3}$$

The following equation is obtained by reducing the normalized variables $\overline{y_k}, \overline{x_k}\,(k = 1, 2, 3)$ to $y_k, x_k\,(k = 1, 2, 3)$, respectively:

$$y_1 = 47.97 - 0.07x_1 - 0.87x_2 - 0.99x_3,$$
$$y_2 = 612.71 - 0.39x_1 - 10.2477x_2 - 0.71x_3, \tag{4}$$
$$y_3 = 183.949 - 0.13x_1 - 2.49x_2 - 0.01x_3.$$

4. RBF Radial Neural Network Prediction Model

From a structural point of view, the RBF nerve network consists of three layers: an input layer, an implicit layer, and an output layer. The number of nodes in this layer is determined by the actual number of input variables; the number of nodes in the implicit layer is difficult to determine and is determined by the complexity of the problem being described [23]; the output layer is used to output the results of the input and output processing of the system. In the RBF neural network, the transformation from the input space to the hidden space is a nonlinear transformation, and the transformation from the hidden layer to the input layer is a linear transformation. Assuming that the RBF neural network has only one hidden unit [24], the training sample of the RBF neural network can be expressed as $\{X_n, d_n\}\,(n = 1, 2, \cdots, N)$. In this formula, $X_n = [x_{n1}, x_{n2}, \cdots, x_{nM}]^{\mathrm{T}}\,(n = 1, 2, \cdots, N)$ is the sample input; $d_n\,(n = 1, 2, \cdots, N)$ is the desired sample output, and the corresponding actual output is $Y_n\,(n = 1, 2, \cdots, N)$; the radial basis function $\varphi(X, t_i)$ is the output of the ith hidden unit, and $[t_{i1}, t_{i2}, \cdots, t_{im}, \cdots, t_{iM}]^{(i=1,2,\cdots,I)}$ is the centre of the basis function; $w_i\,(i = 1, 2, \cdots, I)$ is the connection weight between the ith hidden unit and the output unit.

The topology of a specific single-output RBF network is shown in Figure 1.

When the network input training sample is X_n, the actual output of the network can be expressed by the following equation:

$$Y(X_n) = \sum_{i=1}^{\varepsilon} w_i \varphi(X, t_i). \tag{5}$$

Currently, there are many RBF functions in the field of RBF neural networks, the most common of which are the Gaussian function, the multiple quadratic function, and the thin-slab spline function.

The purpose of data normalisation is to transform all data to values in the range [0, 1] to make subsequent calculations more convenient [25]. The data normalisation formula is

$$X = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \tag{6}$$

The network parameters are updated by the following equation:

$$v_i = v_{i-1} + K_i e_i,$$
$$K_i(zxn_y) = P_{i-1}B_i\left[R_i + B_i^T P_{i-1}B_i\right]^{-1}, \tag{7}$$
$$P_i = \left[I_{zxz} - K_i B_i^T\right]P_{i-1} + q_0 I_{zxz}.$$

In the above equation, $v_i = [w_0^i, w_1^i, c_1^i, \delta_1^i, \cdots, w_K^i, c_K^i, \sigma_K^i]^{\mathrm{T}}$ represents the state of the parameter shown for the ith sample after it has entered the network.
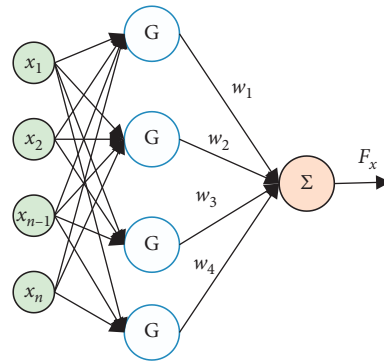
FIGURE 1: Topology of a single-output RBF network.
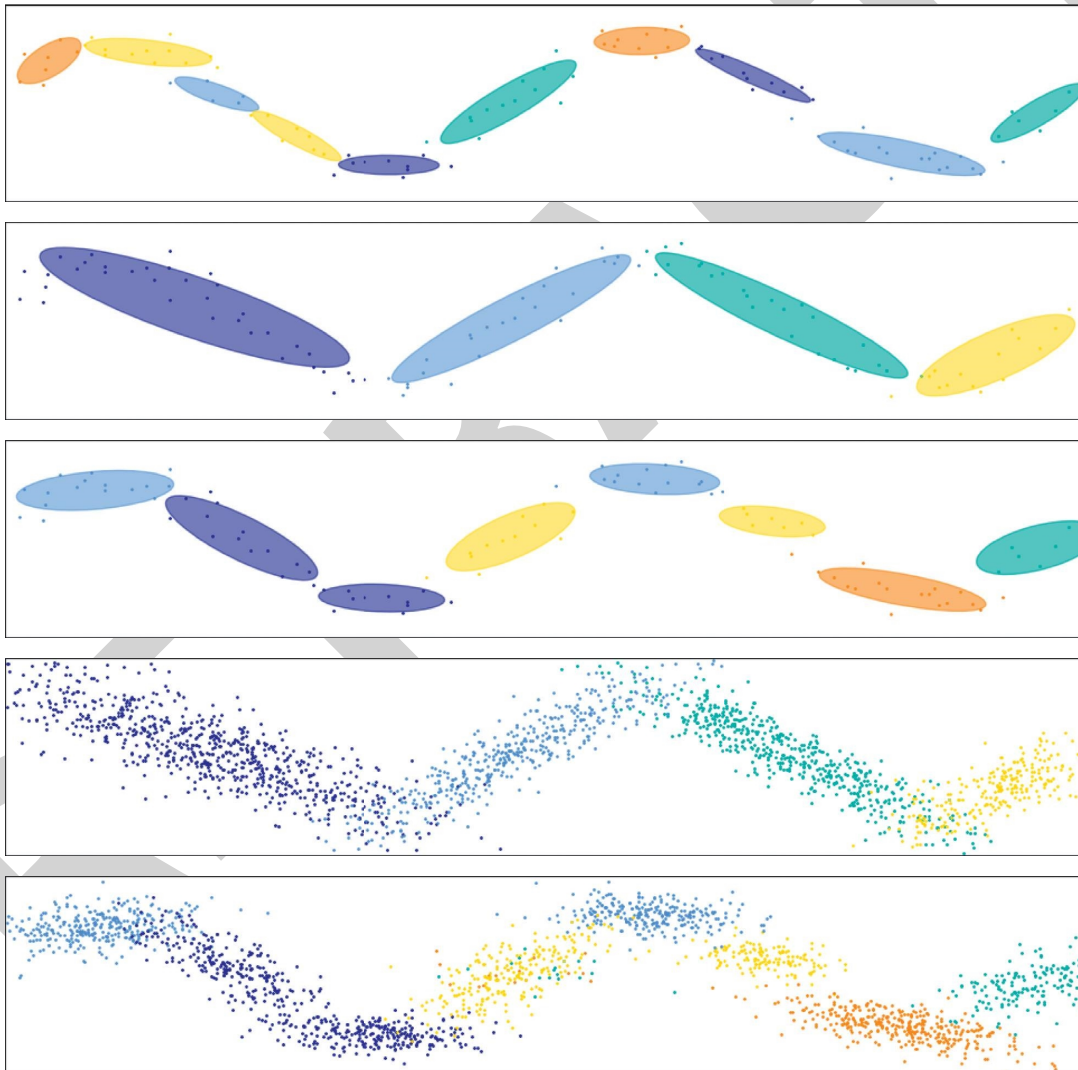
FIGURE 2: Typical training sample of the head state detection classifier.

The calculation is carried out for all hidden node outputs $\sigma_k = (o_{k1}^i, \cdots, o_{kj}^i, \cdots, o_{kn_y}^i)$, $i = 1 \cdots N$, $k = 1, \cdots, K$, $j = 1, \cdots, n_y$, and $\|o_{j,\max}^i\| = \max(\|o_{1,j}^i\|, \|o_{2,j}^i\|, \cdots, \|o_{K,j}^i\|)$, $(j = 1, \cdots, n_y)$. The specific normalized output vector formula is

$$r_{kj}^i = \frac{\left\|o_{k,j}^i\right\|}{\left\|o_{j,\max}^i\right\|} \ (k = 1, \cdots K). \tag{8}$$

If all $r_{kj}^i < \varepsilon_2$ are established, then the kth hidden node can be deleted.

5. Experimentation and Analysis

In order to verify the effectiveness of the method, we experimented on an Intel Core 2 T6500 PC with 3 G RAM, implemented using VVC++6.0 with the OpenCV vision library programming [26]. To train the FSAMME-based head state classifier, we manually intercepted 800 head images from 100 basketball game videos, 100 images from each class. A typical training sample is shown in Figure 2. In addition, 100 nonhead images were intercepted as a negative example training set, consisting mainly of local stadium background blocks and blocks of other body parts of the players. Figure 3 shows the relationship between the relative size (ratio to sample side length) of the small square area extracted from the training sample and the recognition accuracy. In the training process, we set τ_{accuracy} of the head image detection classifier to 0.9, the learning rate τ_{rate} to 0.1, and τ_{accuracy} of the nonhead image detection classifier to 0.4. We selected another 400 head images as the test set, 80 images of each class, and the experimental results are shown in Table 3.

As can be seen from Table 3, the misjudgements were mainly in the discrimination of two adjacent head poses. This is mainly due to the low resolution of the head image in the basketball game video, the interference of motion blur and noise, and the similarity of the adjacent head postures themselves. From the experimental results, it can be seen that the covariance descriptor, which incorporates a variety of underlying features, is well adapted to the above situation and achieves satisfactory experimental results, with the average recognition rate of various head poses.

In order to verify that the covariance descriptor can improve the accuracy of head state recognition after fusing the original underlying features, we have selected a subset of feature vectors in the section for comparison experiments, respectively, and the experimental results are colour features, gradient features, texture features, colour-gradient features, colour-texture features, gradient-texture features, and colour-gradient-texture features, as shown in Figure 4.

In addition, we also experimented with the method in this article in the case of partial occlusion of the player's head. Test cases were manually occluded from the top, bottom, left, and right directions, with each direction taking up 10%, 20%, 30%, 40%, and 50% of the total head image area, shown in Figure 5.

As can be seen from the graph, the accuracy of head pattern recognition is generally higher when the left and
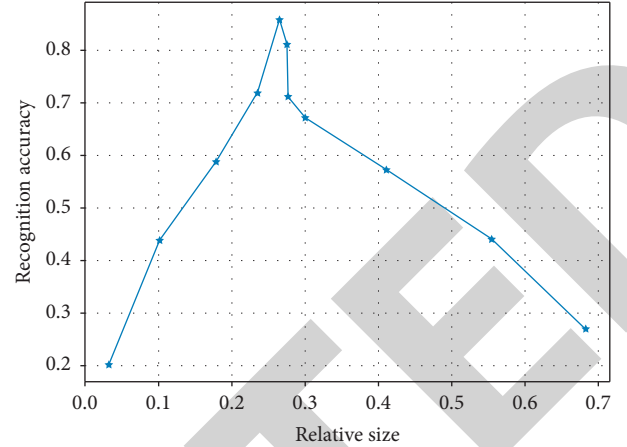
FIGURE 3: Accuracy of head pattern recognition versus relative size of small areas.

TABLE 3: Results of the head state recognition experiment.

Degrees	0	45	90	135	180	225	270	315
0	0.98	0.01	0	0	0	0	0	0.01
45	0.01	0.96	0.02	0	0	0	0	0.01
90	0	0.02	0.95	0.01	0	0.01	0	0.01
135	0	0.01	0.02	0.95	0	0.01	0	0.01
180	0	0	0	0	0.99	0.01	0	0
225	0	0	0	0.01	0	0.97	0.02	0
270	0	0.01	0	0	0	0.01	0.96	0.02
315	0.02	0	0	0	0	0.01	0.01	0.96
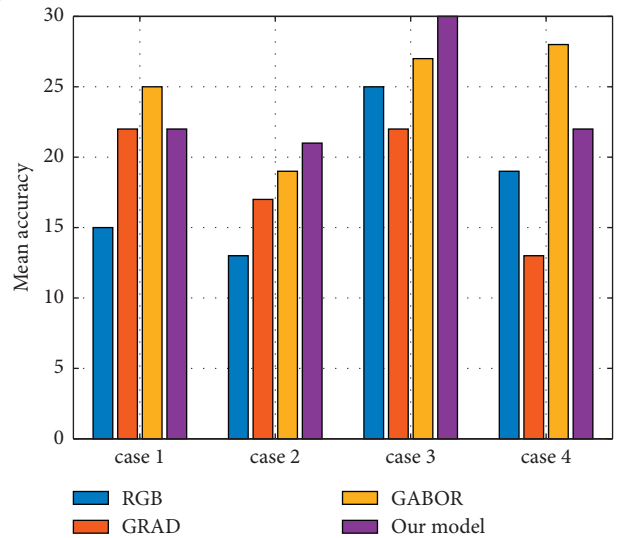
FIGURE 4: Comparison of recognition rates for different types of original feature fusion.

right occlusion occurs than when the top and bottom occlusion occurs. When the occlusion rate is reached, the lowest upper occlusion recognition rate is also higher than the best right occlusion recognition rate by a close margin. It can be seen that the method in this paper also achieves better results when dealing with partial occlusion.
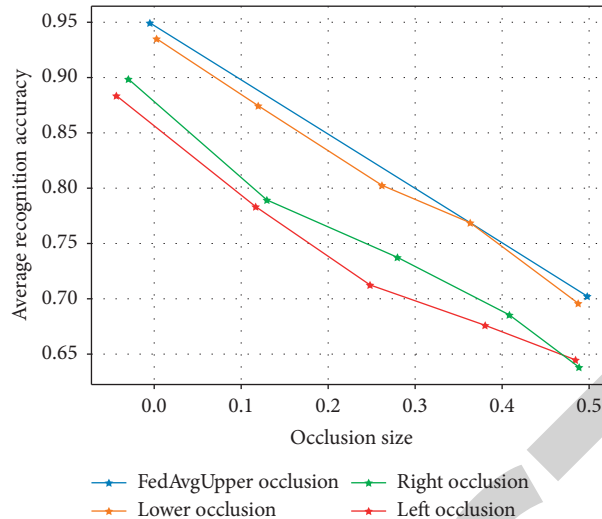
FIGURE 5: Average head state recognition in different occlusion situations.

TABLE 4: Comparison with two existing methods in terms of average recognition rate.

Method	Method in this paper	Orozco	Siriteerakul
Average identification accuracy	96.5%	71.7%	78.6%

Finally, the method is compared with two state-of-the-art methods for low-resolution cephalosporin recognition. As can be seen from Table 4, this method outperforms the other two methods in terms of average head pattern recognition rate.

6. Conclusions

The above analysis shows that the RBF neural network has strong generalization and learning abilities in dealing with the prediction of nonlinear problems, and the above construction shows that the algorithm proposed in this paper has more advantages and accuracy in the prediction of sports behaviour, thus proposing a new scientific training method for the scientific training of current sports.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Q. Feng, "Study on recognition of cross and intermediate features based on RBF Neural Networks," *WSEAS Transactions on Computers*, vol. 5, no. 9, pp. 1831–1836, 2006.

[2] H.-G. Han and J.-F. Qiao, "Prediction of activated sludge bulking based on a self-organizing RBF neural network," *Journal of Process Control*, vol. 22, no. 6, pp. 1103–1112, 2012.

[3] G. Melagraki, A. Afantitis, K. Makridima, H. Sarimveis, and O. Igglessi-Markopoulou, "Prediction of toxicity using a novel RBF neural network training methodology," *Journal of Molecular Modeling*, vol. 12, no. 3, pp. 297–305, 2006.

[4] A. Addeh, A. Khormali, and N. A. Golilarz, "Control chart pattern recognition using RBF neural network with new training algorithm and practical features," *ISA Transactions*, vol. 79, p. S0019057818301666, 2018.

[5] H. Huaien Luo and S. Puthusserypady, "Estimation of the hemodynamic response of fMRI data using RBF neural network," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 8, pp. 1371–1381, 2007.

[6] J. I. Mulero-Martinez, "Analysis of the errors in the modelling of manipulators with Gaussian RBF neural networks," *Neurocomputing*, vol. 72, no. 7-9, pp. 1969–1978, 2009.

[7] K. Z. Mao, "RBF neural network center selection based on Fisher ratio class separability measure," *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1211–1217, 2002.

[8] J. Vales-Alonso, P. López-Matencio, F. J. Gonzalez-Castaño et al., "Ambient intelligence systems for personalized sport training," *Sensors*, vol. 10, no. 3, pp. 2359–2385, 2010.

[9] Q. Song and N. Kasabov, "Twrbf - transductive RBF neural network with weighted data normalization," *Neural Information Processing*, vol. 3316, pp. 633–640, 2004.

[10] Z. Zhenjia, C. Jianning, and A. Jodf, "Study on the method of detecting the crowd abnormality in the sensitive media image based on behavior analysis," *Recent Advances in Electrical & Electronic Engineering*, vol. 9, no. 1, pp. 29–33, 2016.

[11] N. Z. Lima and R. C. Mesquita, "Point interpolation methods based on weakened-weak formulations," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 12, no. 2, pp. 506–523, 2013.

[12] N. Yasynska and O. Ivchenkova, "Using neural networks IN modeling the financial results OF business processes," *WORLD OF FINANCE*, vol. 3, no. 60, pp. 108–120, 2019.

[13] G.-F. Wu, P. Y. Lin, P. J. Huang, M.-Y. Xiao, and J.-F. Zhang, "Application of the Neural Network structural optimization in broiler growth performance prediction," *Guangdong Agricultural Sciences*, vol. 38, no. 22, pp. 126–276, 2011.

Retraction

Retracted: Financial Market Sentiment Prediction Technology and Application Based on Deep Learning Model

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Guo, "Financial Market Sentiment Prediction Technology and Application Based on Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1988396, 10 pages, 2022.

Research Article

Financial Market Sentiment Prediction Technology and Application Based on Deep Learning Model

Yixuan Guo (ORCID)

Department of Business Administration, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China

Correspondence should be addressed to Yixuan Guo; 20099994@lixin.edu.cn

In the real world, there are a variety of situations that require strategy control, that is reinforcement learning, as a method for studying the decision-making and behavioral strategies of intelligence. It has received a lot of research and empirical evidence on its functions and roles and is also a method recognized by scholars. Among them, combining reinforcement learning with sentiment analysis is an important theoretical research direction, but so far there is still relatively little research work about it, and it still has the problems of poor application effect and low accuracy rate. Therefore, in this study, we use the features related to sentiment analysis and deep reinforcement learning and use various algorithms for optimization to deal with the above problems. In this study, a sentiment analysis method incorporating knowledge graphs is designed using the characteristics of the stock trading market. A deep reinforcement learning investment trading strategy algorithm for sentiment analysis combined with knowledge graphs from this study is used in the subsequent experiments. The deep reinforcement learning system combining sentiment analysis and knowledge graph implemented in this study not only analyzes the algorithm from the theoretical aspect but also simulates data from the stock exchange market for experimental comparison and analysis. The experimental results illustrate that the deep reinforcement learning algorithm combining sentiment analysis and knowledge graphs used in this study can achieve better gains than the existing traditional reinforcement learning algorithms and has better practical application value.

1. Introduction

Reinforcement learning is a commonly used framework as a way to handle sequential decision-making tasks. On the other hand, deep learning can be used for the representation and storage of feature aspects [1]. So far, the combination of these two different models is by far the best answer in terms of learning good state representations of very challenging tasks that are designed to solve not only simulated scenario domains but actually challenging real-world problems. Reinforcement learning, as a subset of machine learning, which in turn is an important branch of artificial intelligence, has gained more and more importance in the last years [2]. The classical approach to creating AI requires programmers to manually code every rule that defines the behavior of the software [3]. A clear example is Stockfish, an open-source AI chess engine developed with the help of hundreds of programmers and chess experts who translated their experience into the rules of the game [4]. In

contrast to rule-based AI, machine learning programs develop their behavior by examining large amounts of example data and finding meaningful correlations.

In creating a machine learning-based chess engine, instead of providing rules for each game, engineers created a basic algorithm and trained it using data collected from thousands of games played by human chess players [5]. The AI model would scrutinize the data and identify similarities between winners. When offered a new game, the AI will decide on the moves most likely to lead to a win based on examples it has previously seen. While machine learning and its more advanced subset deep learning can solve many problems previously thought to be infeasible for computers, they rely on large amounts of annotated, quality training data. This limits their application to areas where labeled data are scarce [6]. This is where reinforcement learning comes into play. Humans and higher animals are able to engage in continuous interaction with their external environment to understand, explore, and sense it. This

is due to the fact that humans and higher animals have the ability to continuously learn, to continuously learn and accumulate experience, and to reuse existing experience to make current decisions more rational and superior. The implication of machine learning is that machines have the same learning and perceptual capabilities as humans and can mimic the action habits of humans or higher animals in continuous interaction [7]. Machine learning has a number of advantages over human learning, and it has been shown that machine-based knowledge far exceeds the capabilities of the human brain in terms of memorizing knowledge, understanding, and comprehension. As a result, humans tend to rely more and more on machine-based knowledge. The value of utilizing machine learning in finance is becoming increasingly evident. As banks and other financial institutions strive to enhance security, streamline processes, and improve financial analysis, machine learning is becoming the technology of choice [8]. While it is true that traditional finance is not at the forefront of adopting machine learning, its use in finance is now a hit. It offers new technological services for financial forecasting, customer service, and data security. Profitable stock trading strategies are vital for investment firms [9]. It is used to optimize capital allocation, thus allowing for the maximum utilization of performance, such as expected returns. Earning maximization is based on an estimate of the potential return and risk of the stock. However, it is challenging for analysts to consider all relevant factors in a complex stock market [10].

In the strategy of this study, first, the covariance matrix of expected stock returns and stock prices is calculated. Then, the optimal portfolio allocation is found by maximizing the return for a fixed risk of the portfolio or minimizing the risk for a range of returns. The optimal trading strategy is then extracted by following the optimal portfolio allocation. However, implementing this method can be very complex if the manager wants to modify the decisions made at each time step and consider the transaction costs. The process of stock investment trading can be thought of as a Markov decision process (MDP). Then, we formulate the trading objective as a maximization problem. Considering the stochastic and interactive nature of the trading market, the problem formulation of stock trading, we solve the stock trading problem by considering the stock trading process as a Markov decision process (MDP) and calculate the optimal strategy through dynamic programming ideas. However, the scalability of the model is limited by the huge state space in dealing with the stock market. In this study, however, our work is a study of stock market sentiment analysis and investment strategy algorithms based on deep reinforcement learning, which first uses knowledge mapping techniques to improve the relevance of news headlines, uses them in sentiment analysis to derive the sentiment coefficients of the corresponding stocks for each news item, and applies them to a modified deep recurrent Q-learning network (DRQN) to find the optimal trading strategy in a complex and dynamic stock market.

2. Related Work

In recent years, various methods based on deep reinforcement learning have been used to deal with various types of decision problems. Although the scheme of using deep learning and reinforcement learning in unison has been used in the past, deep reinforcement learning actually got a start in terms of both quality and efficiency after DeepMind proposed the DQN (deep Q network) algorithm, which accomplished the result of playing Atari games using only pure images as input and using a training intelligence throughout [11]. DeepMind then published a modified version of the previously proposed DQG in nature, which was noticed by the academic community, and it is because of this opportunity that deep reinforcement learning has gradually become a cutting-edge research in the direction of deep learning. Most of these reinforcement learning methods can be further divided into model-free-based reinforcement learning and model-based reinforcement learning [12]. In model-free-based reinforcement learning algorithms, the algorithms are able to be applied in several ways. Since it is not necessary to construct a model, all decision-making studies of intelligence are obtained by interacting with the environment [13]. Among the representative methods are Q-learning and some of its improved versions, the Monte Carlo method.

Without a model, the intelligence needs to explore the external environment by interacting with the system with information all the time and training the model with a lot of trial-and-error experience in the process [14]. This is actually one of the drawbacks of model-free reinforcement learning algorithms: the efficiency of using data for simulation is too low. For these reasons, model-free-based reinforcement learning algorithms usually require millions of explorations and environments. With the above analysis, model-free-based reinforcement learning methods are actually like violent exploration, but model-free-based reinforcement learning methods are guided by the strategy of violent search [15]. Moreover, model-free-based reinforcement learning methods tend to be poorly robust, and especially after the external market environment and training work tasks change and evolve, the intelligence can generally ask us to retrain and explore [16]. Model-free methods can directly learn from experience, meaning that they can perform actions in the real world (e.g., robots) or in computers (e.g., games). They then collect rewards (either positive or negative) from the environment and update their value function. This is the main difference with model-based approaches.

The model-free approach works in the real environment for learning. Instead, the model-based algorithm reduces the interaction with the real environment during the learning phase [17–19]. The goal is to build models based on these interactions with the environment and then use that model to simulate other events. Rather than in the actual environment, it is applied to construct the model and obtain the results returned by that model. As mentioned earlier, this has the advantage of accelerated learning, as there is no need to wait for the environment to respond or to reset the environment to some state to resume learning. However, the downside is that if the model is not accurate, we risk completely learning something different from reality. Learning a model involves performing operations in a real environment and collecting feedback [20, 21]. We call this

experience. Thus, for each state and action, the environment provides new states and rewards. Based on these experiences, we try to derive models, which are nothing but supervised learning problems. There are also methods that learn models with the help of neural networks and optimize reinforcement learning methods based on models using traditional theories of optimal trajectories. Such algorithms include those proposed by Levine at Berkeley and Abbeel at OpenAI. Creating a probabilistic model of the system, using a Gaussian process or probabilistic neural network to learn the model, and using the probabilistic model in planning are approaches such as the PILCO family of algorithms, mainly proposed by Rasmussen et al. in the Machine Learning Laboratory at Cambridge University.

In the model-based reinforcement learning methods, if we can define our own cost function, we can apply the model to compute the optimal operation. Control theory has a strong influence on model-based RL [22]. Recent advances in model-free reinforcement learning demonstrate the ability of rich value function approximators to master complex tasks [23]. Dai and Kong pointed out that the multivariate forecasting method can greatly improve the forecasting ability of bond yield [24]. However, these model-free methods require access to an impractically large number of training interactions for most real-world problems. In contrast, model-based reinforcement learning methods can use learned models to quickly achieve near-optimal control under rather restricted classes of dynamics. In environments with nonlinear dynamics, the complex dynamics in model-based reinforcement learning methods require high-volume models, which in turn tend to be overfitted [25–28]. The two forms of reinforcement learning methods have distinct advantages and disadvantages: the expression value estimation model-free methods achieve good asymptotic performance but poor sample complexity, while the modeled methods exhibit learning with good results but struggle with complex tasks.

3. Knowledge Graph-Based Sentiment Analysis of the Stock Market

3.1. System Design. The framework of the knowledge graph-based sentiment analysis system for the stock investment market implemented in this study contains three main modules, knowledge graph module, embedding layer module, and recurrent convolutional neural network (RCNN) module. Relevant news headlines are identified at the top of the website and preprocessed using natural language processing techniques such as removing HTML tags, marking sentences, removing banned words, and so on. Subsequently, the knowledge graph is built by putting the relevant information of the stock corresponding companies, and the preprocessed news headlines are put into the knowledge graph for distance comparison. The news headlines with appropriate distances are filtered in the knowledge graph as the input of the embedding layer module, and then, the sentiment results are output through the sentiment analysis module after vectorizing the text data

in the embedding layer. The specific algorithm flow is structured in Figure 1 below.

The technical development routes for constructing knowledge graphs can be generally divided into two types, one of which is top-down while the other is bottom-up. One of the top-down solutions for constructing a knowledge graph is to find information such as ontologies and relationships from inside the good data and append it to the knowledge base with the help of structured data sources such as encyclopedias. The bottom-up approach to build the knowledge graph is to use specific guidelines to collect data from public resource models to obtain highly reliable information to add to the base knowledge. The knowledge graph constructed in this chapter is based on the cloning of web resources as data input, and therefore, a bottom-up approach to constructing a knowledge graph is used. The input data may be structured, unstructured, or semi-structured, and based on this information data we can apply mainly our own set of systematic automated or semi-automated technical research solutions to obtain the elements of relevant knowledge, i.e., the physical-economic relationships with each other, from the very first data, and then place them in the module layer above the corresponding knowledge base and in the data layer. The construction of the knowledge graph as a process of cyclic change changes by obtaining the relationships of knowledge, and each iteration contains three stages: (1) knowledge extraction: the mutual relationships that exist inside the entities, attributes, and entities are obtained in each data source, and the dynamic knowledge representation is formed based on this logic; (2) knowledge integration: in order to obtain new knowledge after eliminating contradictions and ambiguities such as some entities can be more than one expression, there exists the possibility of some special calls corresponding to a different number of entities; (3) knowledge processing: quality assessment of new knowledge can be placed after their integration before the ranking part of the knowledge base to ensure the quality of the knowledge base. There are two possible problems with the use of remote supervision in relational extraction. First, the knowledge base and the text will use heuristic alignment in remote supervision, which may lead to errors in the annotation of the data. Second, in previous methods, specific features are used in statistical models to achieve classification, but the noise caused by specific features sometimes produces bad results. In this study, multiple example learning and PCNNs are applied to deal with the above two problems. Figure 2 shows the structure of PCNNs.

3.2. Market Sentiment Model Construction. After determining the direction and class of the relationship, we represent the relationship as a form of three tuples (h, r, t). h and t represent the head and tail entities (firms), respectively, and r represents the relationship between the head and tail entities. For example, (Company A, AG, Company B) indicates that Company A holds Company B. After placing the triples into the Neo4j database, the construction of the knowledge graph is completed. The use of vector form to
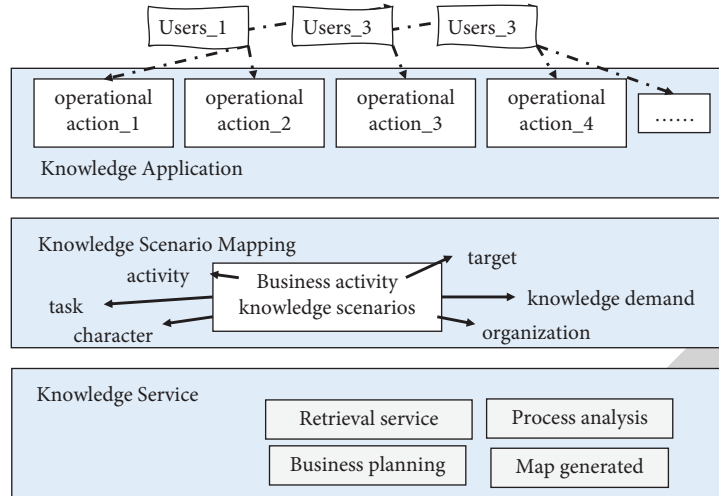
FIGURE 1: Flowchart of knowledge graph-based sentiment analysis algorithm.
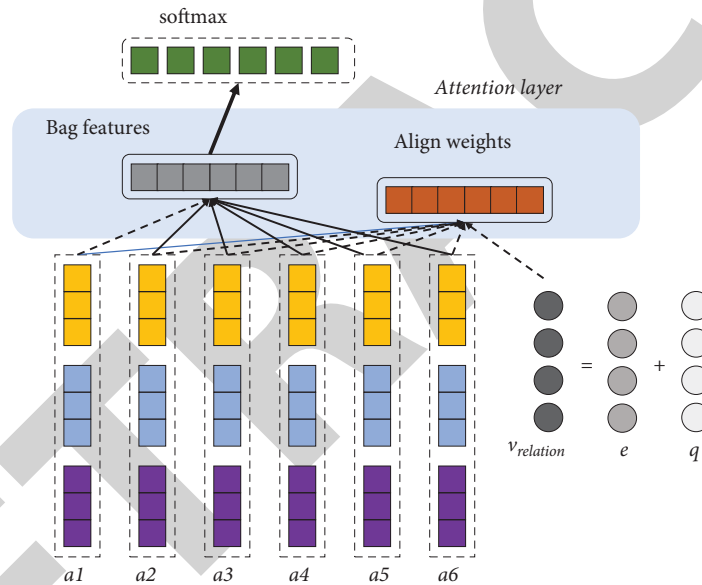
FIGURE 2: Structure of PCNNs.

represent the knowledge graph can be developed to make it more convenient for various social workers afterward; therefore, the goal now is to encode and compress the triad (h, r, t) into a low-dimensional distributed vector. After completing the construction of the knowledge graph, if new news data are directly added to the knowledge graph for relevance analysis, the ordinary Euclidean distance is difficult to meet the processing of various application requirements. In recent years, trans-series models have been widely used for knowledge representation learning of knowledge graphs. Compared with traditional training methods, trans-series models have simple parameters, are easy to train, and are less prone to these problems of overfitting.

Simply put, the relationship inside the triple instance (h, r, t) is compared to the interpretation from entity h to entity t. The implementation keeps adjusting h, r, and t (vectors of h, r, t) so that ($h + r$) is as close as possible to t, i.e., $h + r = t$.

The knowledge graph, trained with feature vectors, can better represent the relationship between knowledge from different graph databases. The news headline data used in this chapter can calculate the distance between the knowledge by inputting it into the knowledge graph for similarity comparison, which can be limited to a fixed distance as a metric for finding implicit relationships between stock companies. A distance smaller than this indicates that this news headline has a closer relationship with the corresponding stock, and there will be factors that affect the rise or fall of the stock, which can be used as input to the sentiment analysis module. News headlines that exceed this distance can be considered as having noisy data and are not used as input for the sentiment analysis module.

The word vector output from the text data after the embedding layer can be used as the input to the sentiment analysis module. The sentiment analysis model network in

this chapter has three layers, i.e., LSTM (recursive) layer, convolutional layer, and output layer. The first layer is the RNN implemented as (LSTM). The LSTM layer introduces the stored elements into the network. This layer is very effective in extracting sentence representations, allowing our model to analyze long sentences. The input gate directs the amount of change in the input vector of the current time LSTM unit to the messages in the memory model, the forgetting gate acts as the degree of influence of the previous moment's history on the messages in the current time memory unit, and the output gate controls the amount of output of the messages in the memory unit. The input gate, output gate, and forgetting gate are denoted as it, ot, and ft below.

The second layer creates a convolution kernel that convolves with the input in a single spatial dimension to generate a tensor. This layer is used for a one-dimensional maximum pool of pool length four. This layer extracts semantic information from the word embeddings provided by the embedding layer. Where said convolution window has a large interference of size etc., similar to the selection to achieve the final classification result of the selected local feature n words, then the representation in the contextual semantic information, i.e., full-text semantic embodiment, is obtained. A convolutional operation window is set to carry out a size of $a * b$, the number of words within the window is a, and the word vector space dimension is b. Multiple convolutional kernels are applied to generate a feature map, then, a maximum pooling layer operation is implemented, and finally, softmax is run to perform the classification.

3.3. Deep Learning for the Decision Process.

Markov decision process, or MDP for short, has been used by a large number of researchers as a general architecture used to solve most reinforcement learning problems and to initiate subsequent theoretical research. The Markov decision process is explained by means of a 5-tuple (S, A, P, R, and γ), where Z is the set of states, which is finite in number, B is the set of actions, which is also is finite, P is the state transfer probability from one current state to the next, R is the reward function in the network, and γ is the discount rate for each cumulative calculation that can be used to calculate the cumulative reward, as shown in Figure 3.

The state transfer probability of the MDP is

$$\widetilde{A}_{x_{II}} = \sum_{i=1}^{q} \chi_{ii} \Leftrightarrow \mu_A(x) = 0. \tag{1}$$

The meaning refers to the map from the operational state of the strategy to the notation used to represent the strategy π, which refers to the chosen state when Z, in the presence of the set of assigned actions, can be expressed as follows:

$$\widetilde{A} = \widetilde{B} = \sum_{i=1}^{q} \mu_{ii}(x) \Leftrightarrow \mu_A(x) = 0. \tag{2}$$

This system is fully visible inside the MDP, which means that the observation is in the same state as the environment: $\omega t = st$. At each time step t, the probability of moving to $zt + 1$ is given by the state transition function $T(zt, bt, zt + 1)$, and the reward is given by the bounded reward function $R(zt, bt, zt + 1) \in \mathcal{R}$. At each step, the intelligence takes action to change its state in the environment and provide a reward.

$$W_{i+1} = W_i - a\frac{\partial L}{\partial q}. \tag{3}$$

The specific meaning of decision-making is available by learning conditions for the probability distribution. Decision-making in reinforcement learning is generally stochastic, and the benefit of using stochastic decisions is the ability to integrate search into the sampling step. Search here is where the intelligence goes through other actions to arrive at a better decision. In practice, various kinds of noise are prevalent, which tend to obey a normal distribution curve, and how to reduce this noise also requires the use of probabilistic aspects of the theory. The goal of reinforcement learning is to find the best decision, and through this best decision found to continue the search for the corresponding best reward, where the best is the one that maximizes the total reward in the end. The cumulative reward is defined.

$$V_{\text{map}} = \frac{\{a_1, a_2, a_3, \ldots, a_n\}}{\max P(X, Y)}. \tag{4}$$

Under different decisions and environments, the value of G is uncertain. Since it is simulated under a stochastic decision π, the cumulative reward is also stochastic. However, the reward E is a value that can be determined by and we are able to perform an interpretation of the state value function. Such a function can represent the state values: when said decision agent is used by π, the cumulative reward can be seen as an allocation, in which E denotes the cumulative value returned by the Z state is in the equation Z value.

$$V_{\text{np}} = \frac{\arg v_j \{a_1, a_2, a_3, \ldots, a_n\}}{\prod P(X, Y).} \tag{5}$$

The corresponding equation for the $z - b$ value function is

$$P_{x,y} = \frac{N \sum_{i=1}^{n} X * Y}{\sqrt{\sum_{i=1}^{n} (X + Y)}}. \tag{6}$$

The Bellman equation for the z-valued function and the z-b-valued function are given below. From the defining equation of the z-valued function, it can be obtained that the Bellman equation for z can be expressed as

$$V_{\text{map}} = \frac{\{a_1, a_2, a_3, \ldots, a_n\}}{\max P(X, Y)} = P(X, Y \mid v_j)p(v_j). \tag{7}$$

4. Experiments and Result Analysis

4.1. Experimental Data and Preprocessing.

Reinforcement learning is really about finding the best policy by simulating the Markov decision process of the environment and then deriving the optimal reward over the best policy by
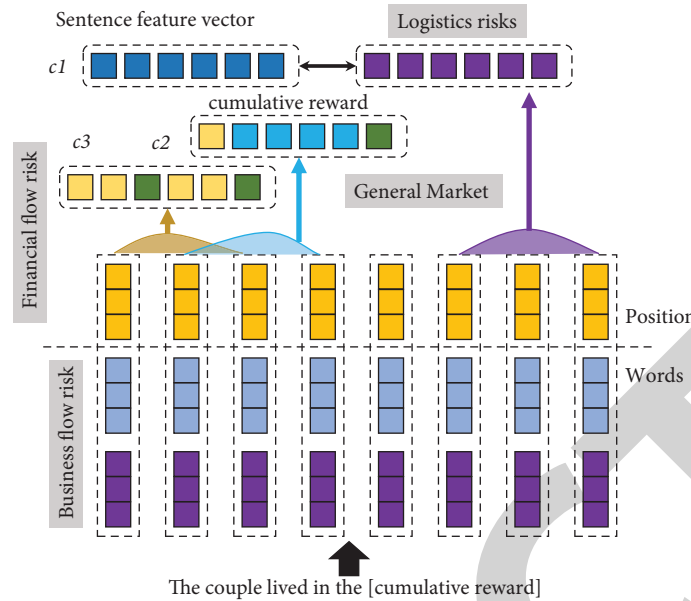
Figure 3: Calculation of cumulative reward.

simulation. To evaluate the effectiveness of the algorithms in this chapter, a large amount of historical data on thousands of stock codes, with both outdated price information and outdated news headlines related to stock codes, is accessible using an interface provided in the financial section. This provides a rich corpus of data that can be used to train agents to correlate their interpretation of headlines with actual changes in stock prices. The site's historical data can be mined and used to train the model. It can then also be examined on a daily basis to make predictions to make predictions. Next, for each news headline, we remove the deactivated words and tokenize them. Each token is then checked for the presence of an organizational node relationship with the particular company of interest in a knowledge graph within a predetermined distance. In the experiments in this chapter, the size of the distance measure chosen is important. Choosing a walking distance longer than this would result in too much noise, while any shorter walking distance would mean that few implicit relationships would be found. The value was empirically adjusted based on conducting some manual experiments. Once any marker in the title was found to be within a predetermined distance of the stock-related organization, by extension, the entire title could be considered relevant to the organization under consideration. In this experiment, 9,463 news headlines for 100 different ticker symbols and 26,017 daily quote instances within the same time frame were collected. The raw HTML data were parsed and filtered to obtain raw tuples of <date, headline> and <date, and quote> instances, which were then sorted by date and merged to provide instances that the model could learn from. As shown in Figure 4, a particular quote instance is classified as RISE or FALL if it is up or down by more than 10% of the previous day's price; otherwise, these data are treated as "None." The prior probabilities of RISE, FALL, or NOTHING in all collected data are 0.014672, 0.008354, and 0.987639, respectively.

4.2. Model Training Setup and Comparison. Because the model has many variable hyperparameters and it is impractical to experiment with the optimal parameters one by one in huge parameter space, the experiments in this section use a grid search approach to perform parameter optimization. To reduce the chance of tuning the parameters for training, the experiments refer to the three-time cross-validation method. For the training process, 80% of the headlines are randomly sampled from the crawled news headline dataset as the training set, 10% of the headlines are randomly sampled as the validation set in the testing process, and the remaining 10% of the headlines are set as the test set to find the best hyperparameters by comparing the three experiments. In this experiment, 9,463 news headlines for 100 different ticker symbols and 26,017 daily quote instances within the same time frame were collected. The loss function is used for cross-entropy, and the gradient descent method for the training optimization process uses SGD stochastic gradient descent method to find the best learning rate through experiments. The results of a simple control experiment show that the best prediction results are obtained when the learning rate is 0.001. Due to the limitation of space, some of the best parameter results derived from the grid search are directly given here, which are the convolutional kernel size 64, the number of convolutional kernels is 4, the number of LSTM hidden cells is 128, the dropout retention ratio is 0.6, and the knowledge graph ontology distance is 5. On this basis, the optimal parameter comparison results for the word vector dimension are made as shown in Figure 5.

The accuracy of KGRCNN is above 86% in the best combination of hyperparameters. The experiments comparing the model in this study with other methods reveal that the KGRCNN method has a great advantage over the common LR and SVM models, while the RCNN model can classify the stock market sentiment better and simply using a CNN or RNN approach does not affect the results and the
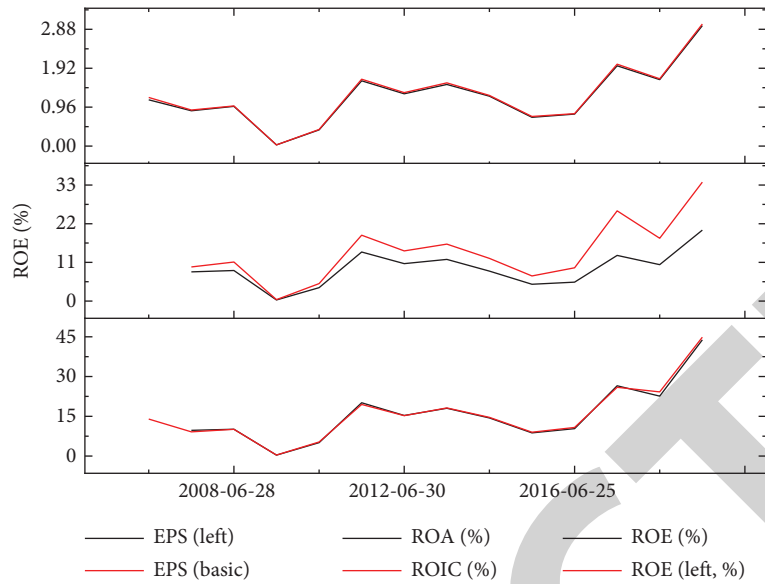
FIGURE 4: Collation of financial data.
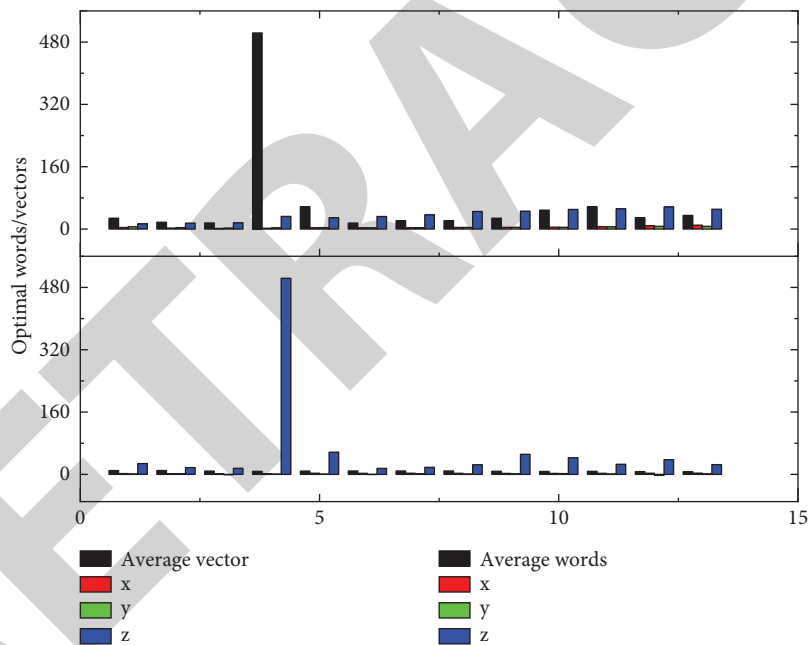
FIGURE 5: Experimental results of tuning the optimal word vector dimension under three cross-validations.

RCNN. A random sample of 90% of the data was used for training, and 10% was reserved for testing. The data chosen for training and testing are randomly selected at runtime. This assumes abstraction of any temporal information that may be present in the data that may affect classification. That is, the training instances were considered independent of each other and only dated so that citations could be matched to titles. The increased amount of training data (of which 10% was reserved for testing and not used for training) was tested against the agents. The experiments were conducted in two of them, the analysis of the prediction accuracy results of the sentiment analysis module without and with the knowledge graph technique.

Experiment 1. First, the first experiment was performed using the full sentiment analysis module, on which the experiments were separately performed on the effect of using or not using the knowledge graph technique on the results of the experiments and the results are shown in Figure 6 below.

Experiment 2. The second experiment uses the knowledge graph technique but removes RNN and CNN from the sentiment analysis module, respectively, to test the effect on the results of the experiment, with the following results in Figure 7. To reduce the chance of tuning the parameters for training, the experiments refer to the three-time cross-validation method.
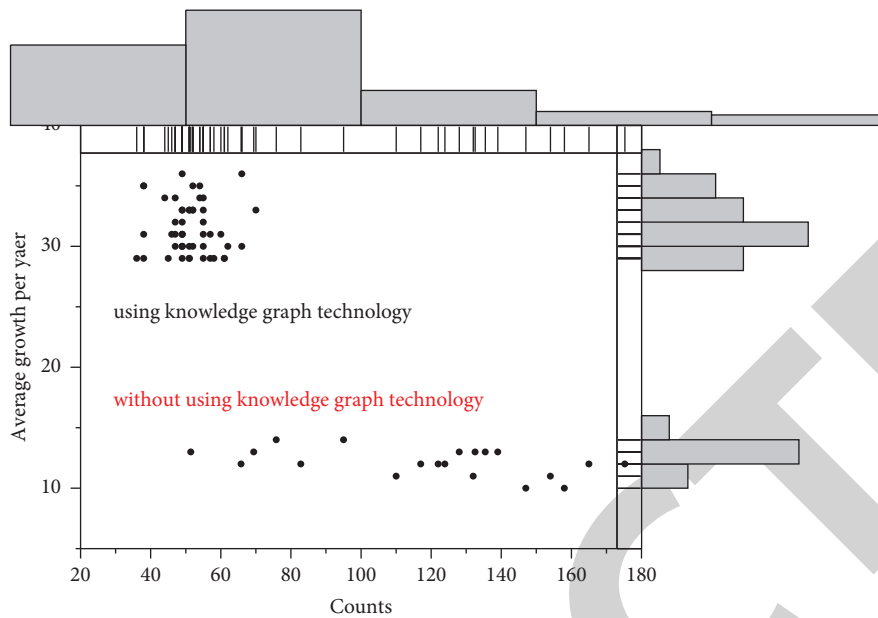
FIGURE 6: Effect of using or not using knowledge graph technology on experimental results.
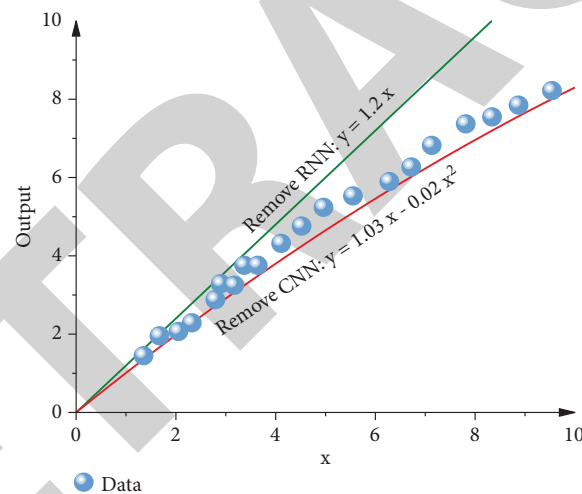
FIGURE 7: Effect of removing RNN and CNN on experimental results.

Experiment 3. The third experiment did not use knowledge graph techniques and removed RNN and CNN from the sentiment analysis module, respectively, to test the effect on the experimental results, which are shown in Figure 8 below.

From the above figure, it can be seen that the combined knowledge graph algorithm used in this study is enough to obtain better accuracy compared to the model without using knowledge graph techniques under the same sentiment analysis module and training data share. The sentiment analysis module using LSTM + CNN also works better than the experiments using only CNN and LSTM. The reason for this is that the internal correlation between real stocks is fully taken into account using the knowledge graph, avoiding the limitation of using only stock price data to measure correlation. RCNN can draw on the advantages of both RNN and CNN. The maximum pool extraction of the convolutional layer extracts the best

representation of the input. The recurrent nature of the network captures a greater degree of pre- and post-textual information when learning text data. The accuracy, recall, F1 scores, and other results of the algorithm steadily improve as the training data increases, and we demonstrate the effectiveness of the algorithm. Based on the idea that knowledge graphs can discover implicit relationships between data, an improved research method for analyzing sentiment problems incorporating relevant knowledge graphs is implemented, and the analysis of the algorithm used in the system reveals that by simply applying a recurrent neural network (RNN) classifier for sentiment analysis, it is unable to identify distinguishing phrases that appear in different orders. In contrast, the convolutional layer can fairly identify the ambiguous phrases in the text through the maximum pooling layer. Thus, the convolutional neural network (CNN) can better reflect the context of the word as
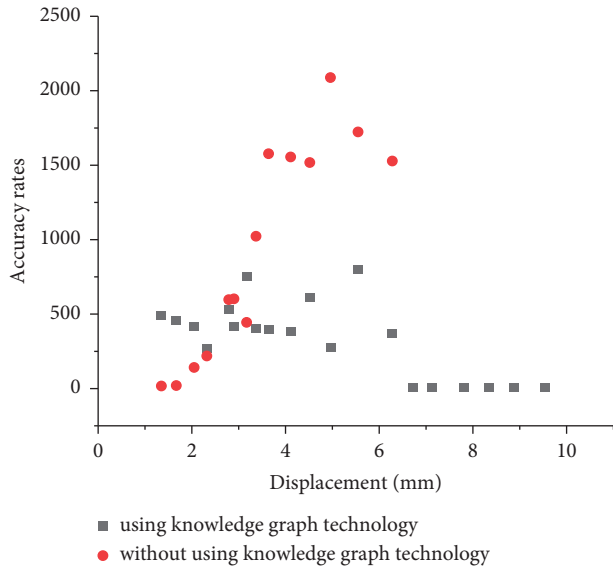
Figure 8: Comparison of accuracy rates in sentiment analysis.

compared to RNN. The system thus chooses recurrent convolutional neural network (RCNN) as the sentiment analysis module for sentiment classification. The maximum pool extraction of the convolutional layer extracts the best representation of the input. The recurrent nature of the network will capture the preceding and following textual information in a greater possible way when learning textual representations. The final experimental results show that when the amount of training data is boosted, the accuracy of the prediction results accordingly increases.

Applying the combination of deep recurrent Q networks (DRQN), prioritized experience replay mechanism, sentiment analysis, and knowledge graph to stock investment strategy research, a deep reinforcement learning algorithm PERDRQN based on deep recurrent Q networks (DRQNs) and prioritized experience replay are proposed to perform stock investment strategy research methods. Simulation experiments using deep learning libraries such as TensorFlow are conducted to verify that the algorithm proposed in this chapter can effectively perform the stock investment strategy task. The algorithm uses an end-to-end training model that combines the feature refinement and combination capabilities of deep recurrent convolutional neural networks and the good strategy formulation capabilities of reinforcement learning, which can obtain information about a stock and its company by crawling data on websites and some publicly available datasets, and after processing it through the knowledge graph and using it as input to the RCNN model, the sentiment of the company's stock is output after network processing, and then, this sentiment analysis combined with the stock information is added inside a deep reinforcement learning model to output an investment strategy. The deep recurrent Q network (DRQN) algorithm combines a convolutional neural network and a recurrent neural network to train the target value function compared to previous Q learning-based reinforcement learning algorithms, optimizing the previous Q learning scheme of training through tables or using linear regression, allowing the algorithm to represent a larger space of action states.

5. Conclusions

The machine has certain intelligence after using deep learning and reinforcement learning. In the process of expanding the theory and practice of machine learning, the theory and results of some disciplines have been integrated, including artificial intelligence, psychology, biology, game theory, and cybernetics, and also interact with many disciplines to get some innovative research directions. Reinforcement learning theory has become one of the hot research directions in the artificial intelligence community after years of pioneering development. However, there are still many problems that need to be improved, such as the balance between exploration and exploitation, the long training period, and the "dimensional disaster." Because the previous work is basically gathered in theoretical research, deep reinforcement learning is relatively little applied in financial investment. In this study, we apply the deep reinforcement learning algorithm PERDRQN, which combines the sentiment analysis method of knowledge graph with the deep reinforcement learning method of prioritized experience replay, based on the study of deep reinforcement learning algorithms for stock investment strategies. Then, the technique of prioritized empirical replay is analyzed and its use in combination with DRQN reduces the time required for convergence and gives better results. The algorithm showed better results in the comparison experiments, indicating that the sentiment analysis approach combined with knowledge graphs is helpful for the deep reinforcement learning strategy framework used in this study. In the future, the accuracy, recall, F1 scores, and other results of the algorithm steadily would be improved as the training data increase.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

Acknowledgments

References

[1] Z. Jin, K. Guo, Yi Sun, L. Lai, and Z. Liao, "The industrial asymmetry of the stock price prediction with investor sentiment: based on the comparison of predictive effects with SVR," *Journal of Forecasting*, vol. 39, no. 7, pp. 1166–1178, 2020.

[2] D. C. Broadstock and D. Zhang, "Social-media and intraday stock returns: the pricing power of sentiment," *Finance Research Letters*, vol. 30, pp. 116–123, 2019.

[3] T. Ito, H. Sakaji, K. Izumi, K. Tsubouchi, and T. Yamashita, "GINN: gradient interpretable neural networks for visualizing

Retraction

Retracted: Promoting Regional Economic Transformation Forecast Based on Intelligent Computing Technology

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Cai, "Promoting Regional Economic Transformation Forecast Based on Intelligent Computing Technology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1835376, 12 pages, 2022.

Research Article

Promoting Regional Economic Transformation Forecast Based on Intelligent Computing Technology

Hongbiao Cai [ID]

Finance Department, Guangzhou Panyu Polytechnic, Guangzhou, Guangdong 511483, China

Correspondence should be addressed to Hongbiao Cai; caihb@gzpyp.edu.cn

The core of the value of artificial intelligence is to integrate with the real economy and become more dependent on local industries. In the entire artificial intelligence + industry development process, many new changes have appeared. This article mainly studies how to promote regional economy based on smart technology. In the context of rapid economic development, the significance of artificial intelligence and the necessity of regional economic transformation are put forward. The model has designed the national science and technology mechanism framework from the four directions of technology guidance, technology service, technology innovation, and technology balance; following the cultivating ideas from products to innovative industries to innovative intelligent environments, due to the low income elasticity of product demand, technological innovation is mainly based on the dissemination of emerging technologies outside the industry and the lack of coordination between companies under the fiercely competitive market structure; the focus is on promoting the continuous improvement of labor productivity. The experimental results prove that the emerging technology industries of enterprises can promote economic transformation in the era of artificial intelligence, provide a reference for better optimization of industrial innovation and development, and provide methodological support for the government to establish sound emerging technology business models and optimized management methods.

1. Introduction

With the continuous development of computer network technology, artificial intelligence has become a product of technology in the new era. At present, there are many artificial intelligence in life, such as smart TVs, smart phones, and so on, which bring convenience and speed to people's lives. It can be seen that artificial intelligence, a modern technology, has an important impact on people's lives. In artificial intelligence, the application of computer network technology plays an important role. The two influence each other and gradually achieve better development. However, the technical equipment of China's traditional manufacturing industry is relatively backward in general, and the penetration rate of intelligence in traditional manufacturing is low. Traditional manufacturing is still under pressure to adjust structure, reduce production capacity, reduce costs, and increase efficiency, all of which make traditional manufacturing intelligent. Upgrading brings great difficulty.

Artificial intelligence is a modern advanced technological productivity. The application of artificial intelligence technology has brought profound changes in production methods and lifestyles, as well as further liberation of the labor force [1, 2]. The artificial intelligence industry will drive the development of the intelligent economy and will also become a new economic growth point in China in the future. Artificial intelligence technology will become a powerful engine for China's national economic and social development [3, 4]. The artificial intelligence industry will promote the real economy to move towards equipment intelligence, structural optimization, and industrial transformation and quality improvement. At the same time, the evaluation index system of emerging technologies affects

traditional industries from the perspective of integrated innovation based on the mechanism of action refines the dimensions of the impact of emerging technologies on traditional industries, which is a further deepening of existing evaluation research and promotes more focused and targeted evaluation analysis. It enriches the perspective of the evaluation research on the impact of emerging technologies on traditional industries.

Foreign research related to financial support to promote economic growth mainly focuses on the relationship between financial growth and economic growth, and considerable progress has been made. Foreign countries use emerging technologies to strengthen the real economy as an important strategic measure and actively promote comprehensive reform and innovation from manufacturers to innovative systems, from business structures to organizational forms, and from development concepts to business models. Goel believes that traditional industries are still playing an important role in promoting the development of new technology industries [5]. Wright proposed that traditional industries can promote and cultivate emerging industries [6]. Arencibia-Jorge et al. use the benign interaction between traditional industries and emerging industries to achieve "two-wheel drive," which is the key to promoting China's economic growth and ensuring sustainable economic development [7].

In terms of domestic industrial structure, due to the modern technological revolution and the rapid growth of productivity, the industrial structure has gradually moved towards modernization. However, traditional industries use primitive production methods for production; that is, they rely on traditional production technologies, and primitive production technologies particularly rely on directly available resources for production, and the use of labor and capital is inefficient. The rise of emerging technologies contributes to the rise of emerging technology industries and economic growth. The optimization and upgrading of Bychkova's traditional industries can make full use of external factors to develop emerging industries [8]. Harris uses examples of high-tech applications in traditional industries to prove that there is a mutually beneficial relationship between high-tech industries and traditional industries [9].

This paper conducts a longitudinal comparative analysis of the influence of emerging technologies on traditional manufacturing steel in the country and horizontally compares the results of traditional equipment manufacturing and the steel industry; on the basis of clarifying the degree of influence of emerging technologies on traditional industries from the perspective of integrated innovation, corresponding countermeasures and suggestions are put forward. At the same time, electronic interviews were conducted with 4 experts, and the data obtained from the interviews were processed and analyzed. The penetration level of emerging technologies into traditional industries reached 0.420, and the level of absorption and absorption of emerging technologies by traditional industries reached 0.512. Therefore, emerging technology industries have the ability to promote regional economic transformation and upgrading very effectively.

2. Research on the Path of Enterprise's Emerging Technology Industry Promoting Regional Economic Transformation and Upgrading

2.1. Internet of Things. Judging from the current situation, the system structure of the Internet of Things is not complete [10, 11]. As we know, the Internet of Things is divided into three layers: perception layer [12], network layer [9], and application layer [10]. Its structural model is shown in Figure 1.

2.2. Smart Environment. The smart environment is a very interactive system. The environment must be able to perceive people's behavior or situation and be able to accurately read the behavior or situation so as to make judgments to provide the people with corresponding intelligent services [13–15]. Among them, self-perceived mobility and intelligent recognition technology are hot and difficult points of current research. Human-friendly built-in hardware devices, convenient two-way communication mechanism, and security and reliability are one of the most basic technologies for building a smart environment [16, 17]. Due to the complexity of computing components in the environment, many differentiation issues have also been brought about in the design of system applications. In order to protect the differences in the bottom layer, deletion and merging must be performed at a higher level. Standard service description, registration, and inquiry complete a series of customized automated processes at the top of the call and the call, reflecting the functions of the intelligent environment [18, 19].

Smart environment is a new idea and new technology that people put forward in order to make it more convenient for them to serve themselves. Not only must the technical factors be checked, but also personal, cultural, moral, social, and legal factors must be further checked to create a truly intelligent environment [20]. Therefore, the smart environment includes many related industries, including distributed intelligence and data exchange. Distributed intelligence means that ASON realizes the distributed intelligence of the network, that is, the intelligence of network elements, which is embodied in the realization of network topology discovery, route calculation, automatic link configuration, path management and control, and service protection by relying on network elements and recovery functions. Data exchange refers to the process of establishing a temporary interconnection path for data communication between multiple data terminal devices for any two terminal devices.

2.2.1. Related Technology. Intelligent environment technology is the cornerstone of the knowledge search system created in this article, mainly used to solve the problem of knowledge acquisition and integration. The smart environment is essentially a structured semantic network, stored in the form of graphs. The technical fields related to the
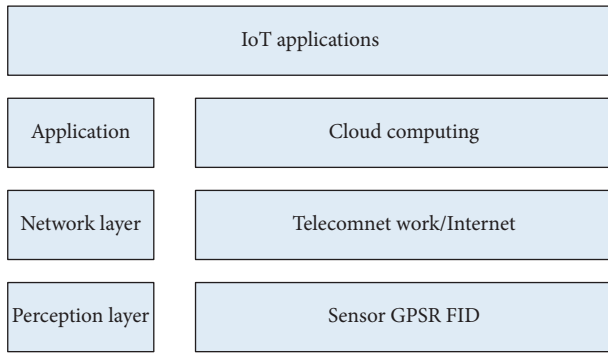
Figure 1: Internet of Things system composition diagram.

smart environment are very extensive [21, 22]. Here, the environment-based computers, technologies related to wireless sensors, and activity recognition technologies are analyzed. These three are key areas in the intelligent environment and are closely related to each other: environment-driven computer environment acquisition, modeling, representation processing, and sensor technology are interrelated and closely related. There are many ways to identify activities, and this article uses a method based on sensor activity. The process of reasoning and recognition activities is the upper and lower processing and combined analysis of the environment [23].

2.2.2. Edge Computing in a Smart Environment.
There are usually two ways to create a smart environment: a top-down approach and a bottom-up approach. The top-down construction method is an ontology-based construction method, using highly structured encyclopedias and other sites as data sources, from which the constraints and rules of the ontology are extracted and supplemented in the knowledge base, while the bottom-up construction method is simple [24, 25]. By identifying patterns and formulating rules, and then adding them to the smart environment, the collected data is used to identify entities, attributes, and relationships. Due to the rapid development of the Internet of Things, the popularity of wireless network technology, and the rapid increase in the number of edge network devices, the data generated by advanced devices has not yet been effectively processed. This is mainly reflected in the linear growth of computing power and massive data. The transmission of network edge devices to the central node leads to a sharp increase in network transmission, which leads to network delays. Network acne data includes privacy, and transmission of raw data to host nodes may leak privacy. In edge computing, data requests come from the local area and are (pre)processed locally and then sent to the remote server. Data processing and transmission are bottom-up methods.

2.2.3. Integrated Smart Environment.
In view of the fact that most research results are based on specific data sources and specific methods, and the built-in intelligent environment of the discipline exposes limited breadth and depth, the universal intelligent environment of the discipline based on bibliographic information design integrates the advantages of various tools and accurately reveals the development of the discipline status quo and the realization of multimethod verification for specific problems [26, 27]. A variety of visualization construction methods integrate a single relationship with data sources, such as keyword co-word relationship clustering tree diagram analysis to reflect the macro structure of the subject, and multidimensional scale analysis to reflect the macro structure of the subject.

2.3. Emerging Technologies Affect the Role of Traditional Industries

2.3.1. The Proliferation of Emerging Technologies.
Dissemination of emerging technologies to traditional industries refers to the provision of emerging technologies to potential technology users through certain channels. This proliferation is mainly due to the technological gap between the two parties and the transformative power of traditional industries. At the same time, many external groups are promoting the spreading process, restrictions, and risk factors. It mainly comes from the technological gap between the two sides and the transformation power of traditional industries. At the same time, the diffusion process is affected by many external promotions, constraints, and risk factors. Only when there is a technological gradient between emerging technologies and traditional industries and there are appropriate external conditions to drive them, emerging technologies will diffuse through the market or other intermediaries and transfer to traditional industries, as shown in Figure 2.

2.3.2. Acceptance of Traditional Industries.
When emerging technologies are transferred to traditional industries through the former market or other intermediaries, the effective penetration of emerging technologies will depend on the acceptance and recognition of emerging technologies by traditional industries. The main body of technology acceptance in traditional industries is enterprises, and the research process of evaluating the impact of emerging technologies on traditional industries from the perspective of specific acceptance of integrated innovation is as follows: enterprise decision makers conduct market demand analysis in combination with the market environment, and on this basis, according to the enterprise's own resource conditions and technical capabilities, it decides whether to adopt and which appropriate emerging technologies to adopt to transform the enterprise, so as to achieve the improvement of the technical level, as shown in Figure 3.

2.3.3. Integration of Emerging Technologies and Traditional Industries.
As the latest technology in the current field, emerging technologies have the characteristics of high penetration. With the penetration of emerging technologies in traditional industries, traditional industries will undergo qualitative changes in terms of technological foundation,

FIGURE 2: Flowchart of the diffusion of emerging technologies to traditional industries.
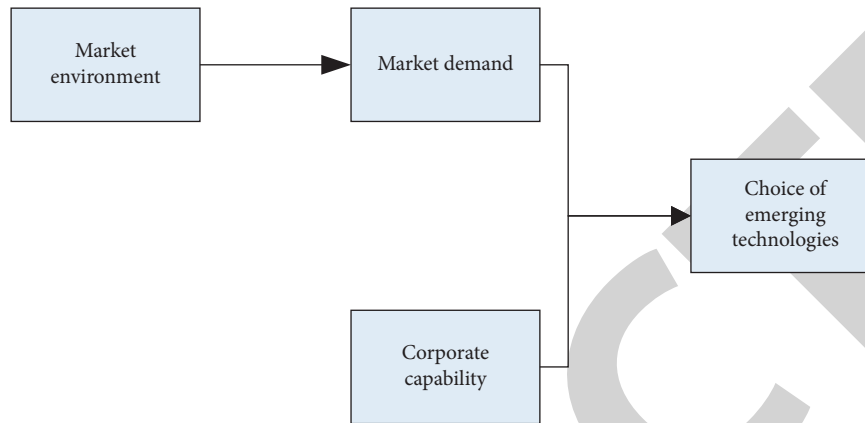
FIGURE 3: Flowchart of traditional industries' acceptance of new and new technologies.

product structure or overall industrial structure. These changes will form interactions in the external environment. The specific fusion process is shown in Figure 4.

As shown in Figure 4, after traditional industries accept the entry of emerging industries, traditional enterprises start from five major aspects, including products, process equipment, and management methods. Promote the improvement of the technological level of enterprises and realize the upgrading of traditional industries.

2.4. The Relationship between Emerging Technologies and Traditional Industries.

The role of integrated innovation is to promote the integration of technological resources and industrial development. The essence of the role of emerging technologies and traditional industries is to accept, integrate, and integrate emerging technologies from traditional industries, thereby promoting the transformation and upgrading of traditional industries. It can be seen that integrated innovation is an important aspect of industrial development [28, 29]. Therefore, this article explores the relationship between emerging technologies and traditional industries from the perspective of integrated innovation. Through the analysis of the tone, stage, and specific impact of emerging technologies affecting the transformation of traditional industries, the specific relationship between emerging technologies affecting traditional industries is discussed from the perspective of integrated innovation, and the theoretical basis of innovation is discussed theoretically.

The impact of emerging technologies on traditional industries is not only the process of technological progress, but also technological and economic activities. On the one hand, the essence of the transformation of traditional industries from emerging technologies is to gradually

upgrade the industrial structure through the diffusion and penetration of emerging technologies, thereby realizing the progress of the industrial structure and leading technological progress [30]; on the other hand, emerging technologies have also provided great help to the transformation and upgrading of traditional industries. The factors of production are being stabilized, as well as the existing economic organizations' understanding of emerging technologies and their participants. This process can effectively improve the production technology level of traditional industries, increase production efficiency, improve corporate governance, effectively reduce product quality, effectively reduce resource consumption and environmental pollution, eliminate useless production capacity, and maintain efficient production capacity. Only when emerging technologies gradually penetrate into traditional industries, and traditional industries decide to accept them, can the two achieve integration and integration, and diffusion and absorption are deeply embedded in the integration process. Therefore, from the perspective of integrated innovation, the essence of the process of emerging technologies affecting traditional industries is the optimal allocation and synergistic integration of existing production factors of traditional industries and innovative factors of emerging technologies.

The introduction of emerging technologies will penetrate all aspects of traditional industries and then lead to a series of changes in traditional industries, such as business philosophy, production management, employment, and systems. It is a process of coordination and integration of various themes or elements, reflecting the characteristics of the system. The impact of technology on traditional industries not only involves the application of new technologies and new processes, but also includes enhancing traditional industries' awareness of concepts, management, markets, and
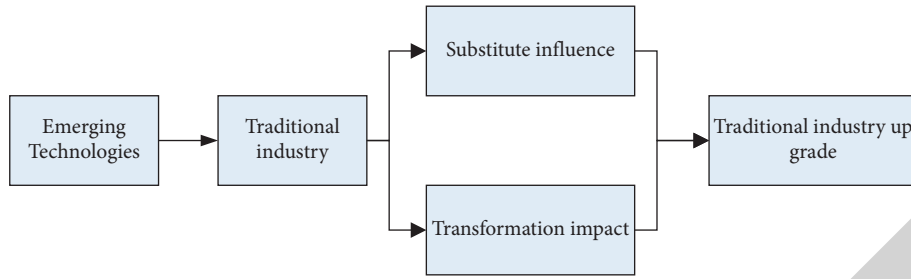
FIGURE 4: Flowchart of integration of emerging technologies and traditional industries.

technological innovation through technology penetration; marketing is one of them. Basically, according to the market environment, it can play its potential from emerging technologies and the sources of emerging technologies and play a role in traditional industries. The impact of emerging technologies on traditional industries is also in line with the basic laws of market economy; the process of emerging technologies operating in traditional industries is a technological and economic activity. The fundamental purpose of using emerging technologies to operate in traditional industries is to promote the transformation and upgrading of traditional industries and optimize production capacity through structural adjustments, thereby stimulating economic growth. It can be regarded as an investment and development behavior, and it is a feature of high technology. In contrast, the biggest feature of emerging technologies is "Xing," which means that they are emerging and coming into effect. In other words, the usefulness of emerging technologies usually takes some time. The use of emerging technologies to transform traditional industries will be a long-term process of change. The absorption of emerging technologies by traditional industries will not only be based on the immediate transformation of emerging technologies, but will also make full use of them. The potential energy of emerging technologies is improving and adapting to the industry.

3. Experimental Research on the Path of Enterprise's Emerging Technology Industry Promoting Regional Economic Transformation and Upgrading in Intelligent Environment

3.1. Test Subject.
In this study, four experts who are most interested in innovation theory or industrial development were invited to evaluate indicators at all levels of the integration and integration of emerging technologies and traditional industries. Experts include 2 X University professors and 2 bosses who have successfully transformed design and manufacturing enterprises. The scoring format is completed through electronic interviews. By constructing a questionnaire about the relationship between evaluation indicators of emerging technologies affecting traditional industries from the perspective of comprehensive innovation, it provides a data basis for analytic hierarchy process used to obtain the weight of each indicator.

3.2. To Gather Data.
According to the expert questionnaire survey, calculate the average value of the expert data, a matrix X can be obtained, which needs to meet $a_{ij} > 0$, $a_{ij} = 1/a_{ji}$, $a_{ii} = 1$, where $i, j = 1, 2, \ldots, n$.

$$X = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}. \tag{1}$$

The square root method is used to measure the weight of each index. The specific formula is shown as follows:

$$h_i = \sqrt[n]{\prod_{i=n}^{n} a_{ij}}. \tag{2}$$

The weights w_i of the possible indicators are as follows:

$$w_i = \frac{h_i}{\sum_{i=1}^{n} h_i}. \tag{3}$$

The calculated indicator weight is only lower than the indicator weight of each standard level. The weight of each indicator in the entire indicator system should be further converted according to the dimensions of the standard level.

3.3. Experiment Procedure

3.3.1. Experiment Preparation Stage.
By reviewing relevant literature, construct a questionnaire on the relationship between the evaluation indicators of emerging technologies affecting traditional industries so that the volunteers can perform the next step in the trial phase.

3.3.2. Experimental Stage.
Interviews were conducted with 4 experts via telephone voice to obtain experimental data. And these data were organized for the next step to calculate the data, and analyze the conclusions of the next step through these results.

3.3.3. End of Experiment.
Discuss and analyze the data results obtained in the experimental stage, combine some knowledge theories and horizontally compare the results of traditional equipment manufacturing and steel industry, and put forward corresponding suggestions on the basis of clarifying the impact of emerging technologies on traditional

industries from the perspective of integrated innovation countermeasures and suggestions.

3.4. Data Processing.
By sorting out the scoring tables filled out by 4 experts, 4 judgment matrices for the integration effect of emerging technologies and traditional industries were calculated. The judgment matrix of the second-level index is u; the judgment matrix of the third-level index is u_1. Since this paper has a three-level index system, it is necessary to carry out a multilevel fuzzy comprehensive evaluation. That is, the first-level comprehensive evaluation is carried out on the first-level scope, and then the second-level comprehensive evaluation is carried out according to the evaluation results, and so on.

$$u = \begin{bmatrix} 1 & 0.9 & 3 \\ 1 & 1 & 5 \\ 0.3 & 0.2 & 1 \end{bmatrix},$$

$$u_1 = \begin{bmatrix} 1 & 0.78 \\ 1/0.78 & 1 \end{bmatrix}. \tag{4}$$

The statistical data used in this article has a different unit dimension for each index data, so the corresponding calculation cannot be directly performed in the overall evaluation. First of all, index data must be processed without dimension, that is, normalization. The common method of data standardization is usually min-max standardization. This method can evenly distribute the index data to the range [0–1]. The indicators selected in this paper have practical significance and are all positive indicators, so it is more suitable to use this method for dimensionlessization. The calculation formula of this method is shown in

$$x = \frac{x - \min}{\max - \min}. \tag{5}$$

The specific formula is as follows:
Large trapezoidal distribution:

$$r(x) = \frac{x - c}{d - c}, \quad c < x < d. \tag{6}$$

Small trapezoidal distribution:

$$r(x) = \frac{b - x}{x - a}, \quad a < x < b. \tag{7}$$

For the ideal score value, it can usually be set as the middle value of each interval. This paper divides the evaluation grades into five grades, and the index data after normalization processing all fall in the [0–1] interval.

3.5. Clustering Algorithm Based on Intelligent Technology.
Clustering originated from data taxonomy, so clustering and classification have a certain connection and similarity, but there is still a certain difference from classification. The clustering algorithm is shown in Figure 5.

As shown in Figure 5, it can be seen that the essence of clustering is not knowing the category or other prior knowledge of each sample in a batch of samples in advance.

3.5.1. Cohesive Hierarchical Clustering Method.
Contrary to the split hierarchical clustering method, the agglomerative hierarchical clustering method first treats each object as a class and then aggregates these nodes in turn to form the nodes of the upper level tree until all the nodes are merged into one large category, or until a termination condition is reached, as shown in Figure 6.

As shown in Figure 6, let the random matrix be W, and let $\lim W^m$ be the primitive random matrix of $W = \lim W^m = \begin{bmatrix} c^\infty & 0 \\ r^\infty & 0 \end{bmatrix}$; then, the following matrix is a stable random matrix as follows:

$$W = \lim W^m = \begin{bmatrix} c^\infty & 0 \\ r^\infty & 0 \end{bmatrix}. \tag{8}$$

Suppose a set of samples is a_i, and the sample has n features; then, the sample vector is

$$a_i = \{a_1, a_2, \ldots a_n\}. \tag{9}$$

Taking Euclidean distance as the distance between the sample and the cluster center, the process of minimizing the transition of the clustering process is given by

$$f(a, b) = \sum_{j=1}^{d} u_{ij}. \tag{10}$$

Finding the smallest point (a, b) in formula (10), Euclidean distance is

$$u_i = \frac{\sum_{j=1}^{d} u_{ij} * a^i}{\sum_{j=1}^{d} u_{ij}}. \tag{11}$$

The cluster center u_{ij} of the class is

$$u_{ij} = \frac{1}{\sum_{i=1}^{c} \left(2/d_{ij}\right)^2} * \left(\frac{2}{d_{ij}}\right)^2. \tag{12}$$

The clustering feature tree is a highly balanced tree that is used to store the feature information of the cluster, and it occupies a small space and can be stored in the memory, thereby improving the clustering speed and scalability of the algorithm on large data sets, as shown in Figure 7.

As shown in Figure 7, the algorithm is mainly divided into two stages: the first stage uses hierarchical clustering; after meeting certain requirements, other division methods are used for macro-clustering processing.

It is assumed that the current optimal position of the entire fish school is to speed up the search speed of the algorithm. In the rear-end collision operation, the global optimal position at this moment is used to replace the optimal position of the individual fish neighborhood.

$$a_i = a_i + \text{rand}(\text{step}) \frac{a_{g\text{best}} - a_i}{\left\| a_{g\text{best}} - a_i \right\|}, \tag{13}$$

where $a_{g\text{best}}$ represents the optimal position of the entire school of fish.
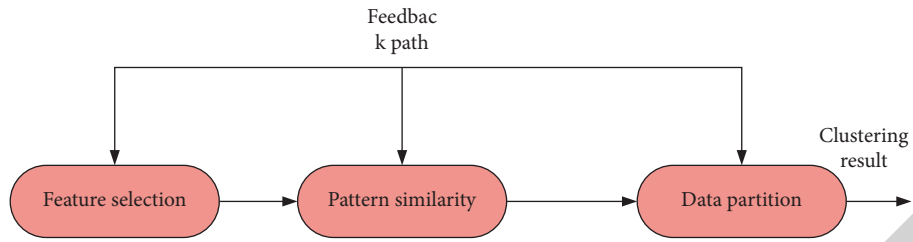
FIGURE 5: Clustering algorithm.
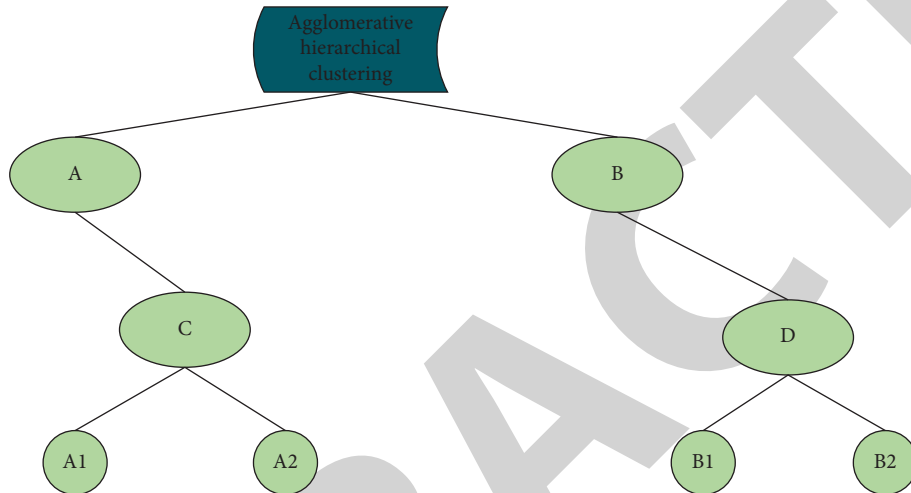
FIGURE 6: Agglomerative hierarchical clustering method.
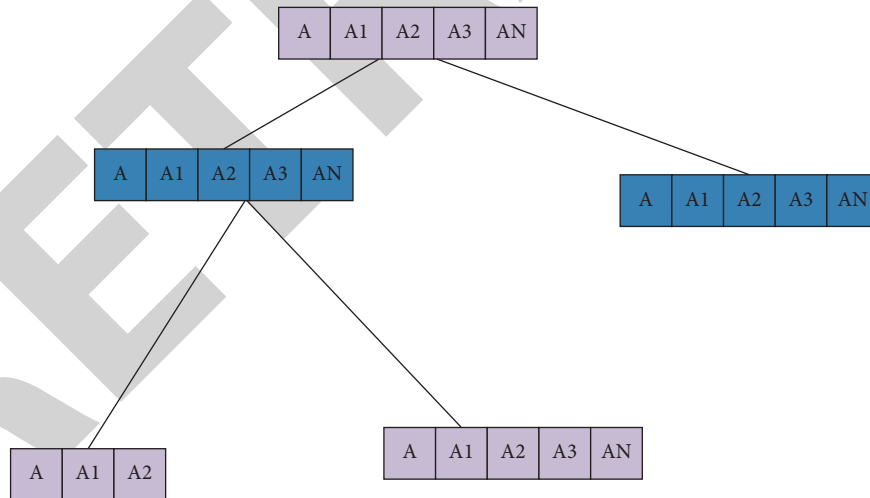
FIGURE 7: Clustering feature tree.

Then, the access sequence length weight can be calculated as

$$lw_i = 1 - \left| \frac{len_i - len}{len_{max} - len} \right|. \tag{14}$$

Taking lw_i as a penalty factor means that the closer the length of the access sequence is to the length of the average access sequence, the larger the value of the factor is, and vice versa, the smaller the value of the factor is.

3.6. *Data Preprocessing Based on Smart Technology.* Preprocessing is the most critical link to ensure that the correct and useful patterns can be discovered efficiently and quickly. The function of this module is to process log files in different types of formats and convert them into data in a unified format that the system can handle; at the same time, it is necessary to complete the cleanup of log data and the identification of access sequences, as shown in Figure 8.
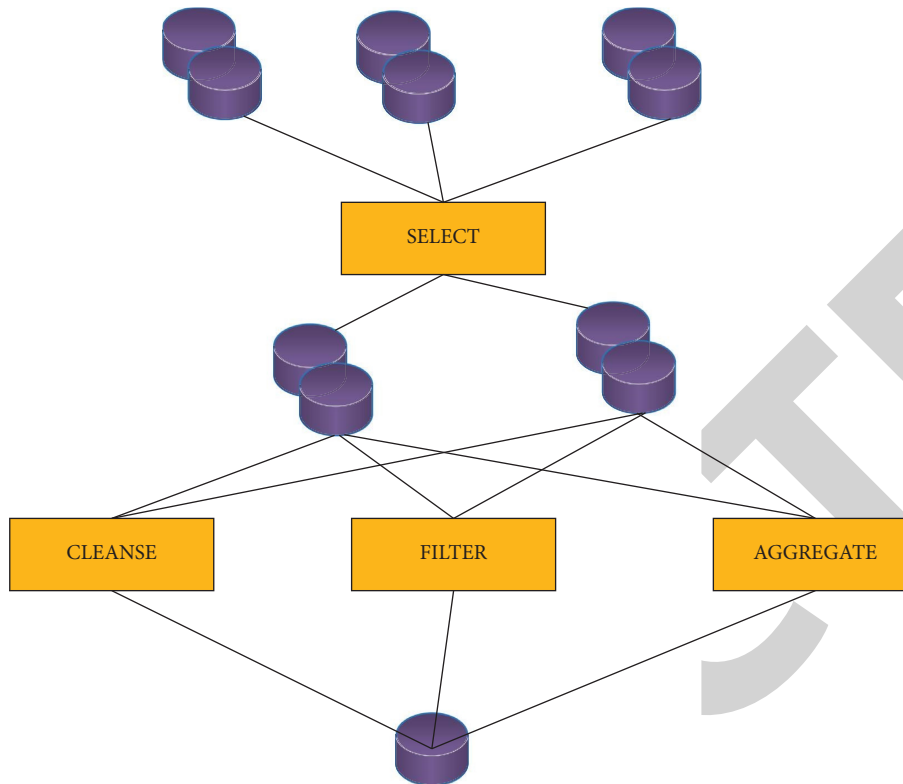
Figure 8: Data preprocessing.

As shown in Figure 8, in this general model of intelligent computing, the key issue is to study how simple individuals can form high-level systems with complex intelligent behaviors under the guidance of simple interaction rules, as shown in Figure 9.

As shown in Figure 9, a complex computing task is completed by a large number of nonsimultaneous computing behaviors of computing units; all the characteristics of the units performing these tasks 3 are unknown to other units and even the system itself; the behavior of a large number of units is determined based on the fact that they are based on incomplete knowledge of the system and delayed or even contradictory information.

4. An Experimental Analysis of the Path of Enterprises' Emerging Technology Industry in Promoting Regional Economic Transformation and Upgrading in an Intelligent Environment

4.1. Analysis of the Weight of Emerging Technology Industry in Promoting the Transformation and Upgrading of Regional Economy. According to the formula of calculating the index weight by the analytic hierarchy process, the weight of the first-level index can be obtained, and the reliability CR = 0.022 is less than 0.05, so the weight distribution meets the requirements. Furthermore, the weight of each secondary indicator can be obtained. The reliability CR = 0.013 is also less than 0.05, and the weight distribution also meets the requirements. Therefore, the weight of each three-level

indicator can be obtained and converted into the overall indicator weight. The impact of emerging technologies on traditional industries is shown. The total weight of the impact evaluation is shown in Table 1.

It can be seen from Table 1 that the data obtained from the experiment is processed and analyzed. The impact level of emerging technologies on traditional industries reaches 0.420, and the external support level is only 0.114, which can make the traditional industries absorb emerging technologies as high as 0.512, so the emerging technology industries are promoted. The transformation and upgrading of the regional economy has been very effective.

4.2. Comparative Analysis of Indicators within and between Industries over the Years

4.2.1. Comparative Analysis within the Industry. Figure 10 shows the impact level of emerging technologies on the traditional equipment manufacturing industry and the steel industry, the level of digestion and absorption of traditional industries, the level of external support, and the effect of integration from 2015 to 2020. On this basis, corresponding countermeasures and suggestions are put forward, as shown in Figure 10.

It can be seen from Figure 5 that, as time goes by, the integration and integration of emerging technologies and traditional equipment manufacturing has continued to increase. First, the penetration level of emerging technologies in the traditional equipment industry is on the rise, reaching the highest level by 2019. The main reason is that the traditional equipment industry has continuously strengthened
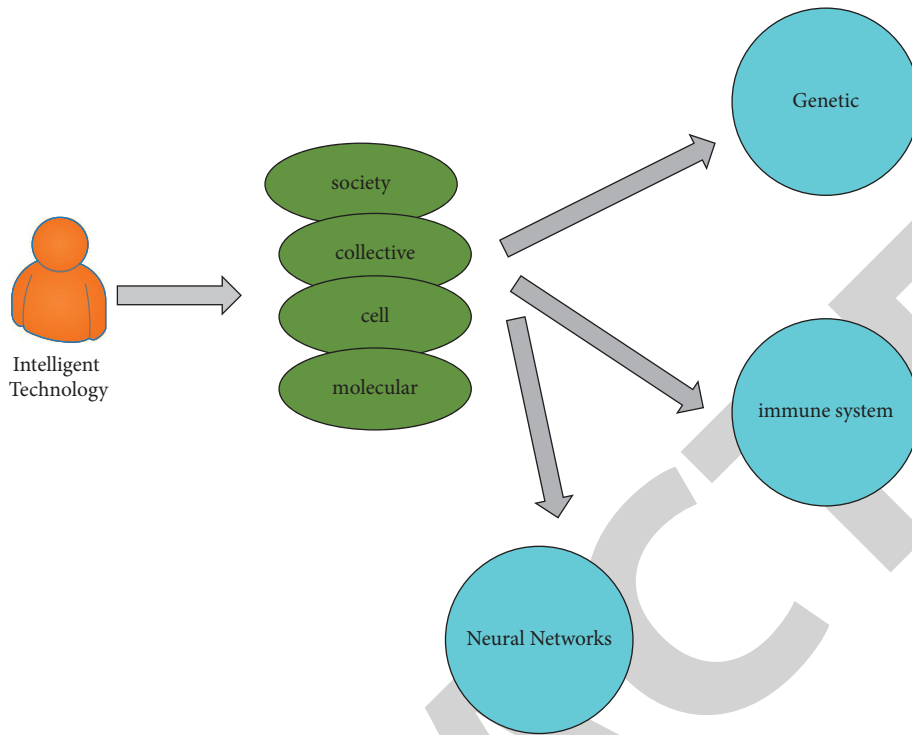
FIGURE 9: Intelligent computing theoretical model framework.

TABLE 1: Weight distribution of evaluation index system.

Level indicator	Secondary indicators	Level-three indicators	Weights
The effect of the integration of emerging technologies and traditional industries	The level of influence of emerging technologies on traditional 0.420	Emerging technology introduction expenditure 0.132	0.372
		Technology and absorption expenditure 0.113	0.238
		Technical transformation expenditure 0.063	0.136
	Traditional industries' digestion level of emerging technologies 0.512	New product sales revenue 0.217	0.367
		The total profit 0.057	0.112
		Number of new product development projects 0.118	0.210
	External support level 0.125	Funds for scientific and technological activities 0.050	0.429
		Tax deduction for technology development 0.057	0.547

the introduction and absorption of new technologies in recent years. Under the influence of new technologies, there is a steady-increase investment in existing technology transformation and R&D.

4.2.2. Comparative Analysis between Industries. Through internal comparison, we can discover various reasons or factors of how emerging technologies in industry effectively affect traditional industries. In order to further analyze the common factors of the impact of emerging technologies on different traditional industries, it is also necessary to conduct a horizontal comparison and analysis of the evaluation results of different industries to find common points and make

more comprehensive recommendations, as shown in Figure 11.

It can be seen from Figure 6 that from the perspective of the penetration level of emerging technologies in the traditional equipment industry and the steel industry, the introduction of emerging technologies in the steel industry was earlier and led the traditional equipment industry before 2015. Since 2015, the traditional equipment industry has been stable. However, the steel industry has experienced a shocking decline. The steel industry is gradually surpassing the traditional equipment manufacturing industry. The traditional equipment industry is more receptive to emerging technologies, and the ability to raise funds from all aspects is becoming stronger.
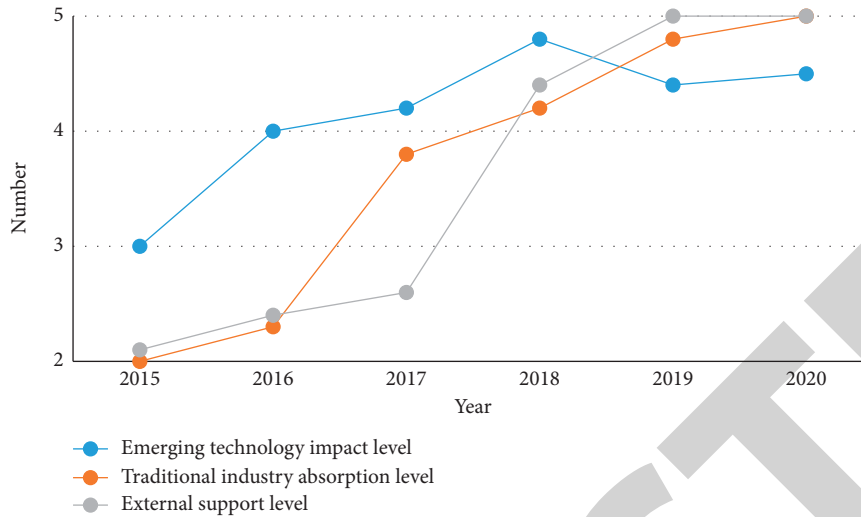
Figure 10: Comparative analysis of various indicators of traditional equipment manufacturing.
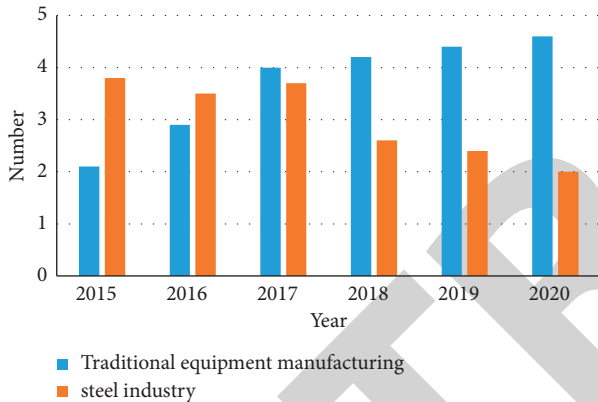
Figure 11: Comparison of penetration levels of emerging technologies in the two major industries.

After a comprehensive analysis of the results, we found that the evaluation results of the integration and fusion of traditional equipment manufacturing and steel industry with emerging technologies are all on the rise, which shows that the integration and integration of emerging technologies and traditional industries is increasing. The bad gradually turned into better or even very good. This also reflects from the side that my country's traditional equipment manufacturing industry and steel industry have effectively used emerging technologies to promote the gradual innovation and development of the industry.

4.3. Analysis of Hausman Test Method

4.3.1. LM Detection Analysis. This document uses the least squares method to estimate the usual table data model, which does not consider the spatial correlation, and then verifies the spatial correlation of the model residuals. The main test methods are Lmsar, Lmerr, R-Lmsar, R-Lmerr,

and so on. If Lmsar is better than Lmerr, and R-Lmsar is important, and Rt-Lmerr is not important, the spatial delay table data model is better than the spatial error table data model. See Figure 12.

According to the results of the LM test, determine whether the spatial measurement model should use fixed effects or random effects. The result of LM test passed the significance test of 0.005. Therefore, this article will deal with fixed-effect time models, fixed-space models, and dual fixed-space and time models. The estimated results are shown in Table 2.

Combining the results of Figure 7 with the results of Table 2, it can be seen that the coefficient of determination and the adjusted coefficient of determination of the spatial panel regression results are significantly better than the ordinary panel regression results, which indicates that the spatial table data model is effective in affecting coupling and coupling. Factors have stronger explanatory power. Commonly used tabular data model analysis industry and traditional industry.

4.3.2. Unit Root Risk Analysis. The test methods are Lmsar, Lmerr, R-Lmsar, and R-Lmerr through natural logarithmic transformation to obtain their ADF statistic values at 1% critical value, 5% critical value, and 10% critical value and perform a unit root test on it as shown in Figure 13.

It is easier to obtain ADF test statistics and the corresponding critical values of Lmsar and Lmerr sequences. When the significance level is 1%, the value of statistical control is greater than the critical value, so the Lmsar and Lmerr sequences are both nonstatic sequences, and both are 1. The level differences of R-Lmsar and R-Lmerr tests are both smaller than the corresponding critical value. It can be seen that R-Lmsar and R-Lmerr are stable. Therefore, it can be determined that the root unit roots of Lmsar and Lmerr are a series of first-order integer monosyllables, satisfying the hypothesis of coupling test.
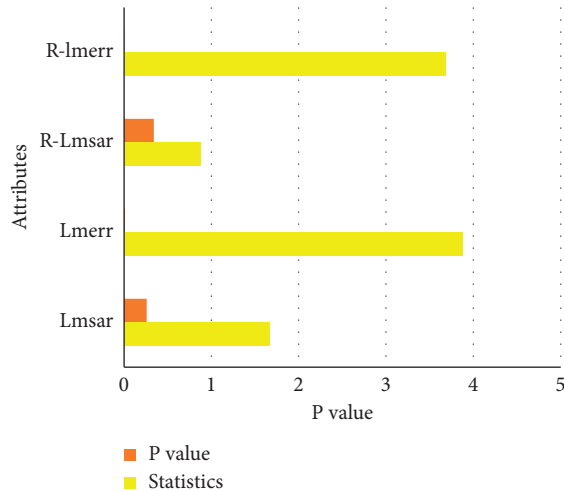
Figure 12: LM test results.

Table 2: The results of the LM test.

Variable	Fixed effect		
	Fixed time	Space fixed	Fixed time and space
Lmsar	0.187	0.176	0.148
Lmerr	3.122	0.469	0.332
R-Lmsar	0.134	0.114	0.085
R-Lmerr	3.117	0.452	0.324
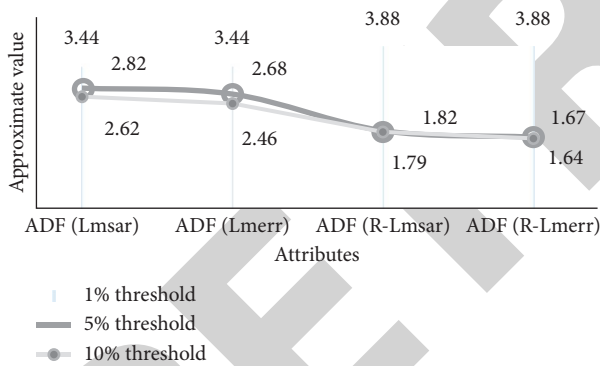
Figure 13: ADF inspection result graph's approximate value.

5. Conclusion

This article mainly expounds the concept of intelligent computing technology and regional economic transformation. In the method part, it proposes a clustering algorithm based on artificial intelligence technology and uses the clustering algorithm to analyze how to play its role in promoting regional economic transformation. As an important subject of development economics, regional economy has strict connotations: how to effectively use human, material, and financial resources to achieve optimal development in a geographical area with basically similar natural environment and humanistic conditions. It is indispensable to constantly update. The research results show that the intelligent computing technology industry has a strong impact on the domestic regional economy in both the short term and the long term, and cities, especially megacities, have further improved the production automation and production. The degree of intelligence, and at the same time, more and more capital replacement labor, has promoted the increase in total factor productivity and promoted the high-quality development of the regional economy; the vigorous development of the domestic regional economy has promoted the sound development of intelligent computing technology, but the intelligent computing technology has further expanded regional economic development differences. Finally, in order to promote intelligent computing technology to promote the high-quality development of regional economy, it is proposed to further create a good innovation environment for the development of intelligent computing technology.

The influence of the integration of emerging technologies with China's traditional arms industry and steel industry is steadily increasing every year. This shows that from the perspective of integrated innovation, the impact of emerging technologies on Chinese traditional industries is increasing every year, which is consistent with the emergence of emerging technologies. Traditional industries have relatively low direct consumption rates and overall consumption rates for strategic emerging industries, while consumption rates for new energy, new materials and latest equipment industries are relatively high. For the new generation of information technology, the consumption rate of organic, energy-saving, and environmentally friendly, and the new energy automobile industry is very low.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest with any financial organizations regarding the material reported in this manuscript.

References

[1] Q. Wang and P. Lu, "Research on application of artificial intelligence in computer network technology," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 5, Article ID 1959015, 2019.

[2] H. Chen, Y. Shang, and K. Sun, "Multiple fault condition recognition of gearbox with sequential hypothesis test," *Mechanical Systems and Signal Processing*, vol. 40, no. 2, pp. 469–482, 2013.

[3] Z. Lv, Y. Han, A. K. Singh, G. Manogaran, and H. Lv, "Trustworthiness in industrial IoT systems based on artificial intelligence," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1496–1504, 2020.

[4] M. Adil, M. K. Khan, M. Jamjoom, and A. Farouk, "MHADBOR: AI-Enabled Administrative Distance Based Opportunistic Load Balancing Scheme for an Agriculture Internet of Things Network." *IEEE Micro*, vol. 42, 2021.

[5] R. Goel, "Indian IT industry sharpens focus on emerging technologies," *Dataquest*, vol. 33, no. 16, pp. 70–72, 2015.

Research Article

User Identity Recognition Based on Wireless Sensor Network and Internet Finance Development

Tianxin Hua ⓘ **and Lingling Zhang**

Jilin Normal University, Siping, Jilin 136000, China

Correspondence should be addressed to Tianxin Hua; huatx@jlnu.edu.cn

With the rapid development of computer network technology, the concept of "Internet +" has become more and more popular in recent years. The combination of the Internet and finance has particularly attracted people's attention, and the operating modes of many industries have also changed. Since the use of Internet technology can achieve data sharing and information exchange, the "Internet + Finance" model has broken the barriers of information asymmetry in the financial sector in the past and has made great contributions to China's multiple improvements. The financial market is very important to China's economic development. The identification of the ID function of the wireless sensor network is susceptible to interference and the identification accuracy is reduced. We propose an adaptive identification feature recognition algorithm based on an improved minimum gray tree. After calculating the similarity, the nearest neighbor matching algorithm is directly used to obtain the minimum matching cost corresponding to the wireless sensor network registration that is regarded as the recognized identity so as to realize the identity function adaptive recognition. In this regard, the simulation results show that the proposed algorithm has high recognition accuracy. With the pace of financial innovation, financial institutions have achieved rapid development on the basis of Internet service platforms. At the same time, as the core of preventing money laundering activities, financial institutions are also facing many issues in identifying "customers" in their work. This article analyzes the main content, implementation effects, and difficulty of customer identification in financial institutions and proposes relevant improvement plans.

1. Introduction

First, the recognition based on the characteristics of the action requires high-precision recognition of the action. With the increasing popularity of wireless object networks and the continuous development of wearable sensors, the value of building sensor-based human behavior and identity recognition systems is getting higher and higher [1]. With the rapid development of information globalization in today's society, the digitization and ambiguity of user ID caused by the Internet will lead to serious disclosure of personal information, which will undoubtedly bring new challenges to identification and technology. The security authentication methods currently in use include combined authentication of user names and passwords, authentication of special authentication objects, and so on [2]. These authentication methods are easy to be stolen or imitated due to

excessive human factors in the authentication process and cannot achieve high-level and high-precision security prevention and control [3]. Due to the convenience of storage such as password authentication and identification technology, users often use passwords and other reasons are relatively simple. There are illegal uses of general username and password combinations, a Bluetooth database has been established, and general IDs have been collected on a large scale. The combination of authenticated username and password has a serious impact on personal privacy and information security [4]. Although it is difficult to avoid the danger due to objective factors, it can be eliminated by the authentication method using biometric properties. Because it is difficult to change the user's biometric characteristics, it is more difficult to imitate, so the possibility of theft is very small. It is now widely used in government, military, banks, welfare companies, social security bureaus, e-commerce, and

other fields [5]. The Internet finance model has been favored and valued by more and more users and enterprises for its advantages of high efficiency, low threshold, convenience, and low cost [6]. It is booming. Various Internet finance platforms are growing exponentially, giving China's financial industry a boost [7]. Development has brought new opportunities. However, the development of new things often has two sides, and the rapid development of Internet financial platforms has led to a saturation trend in the market. Therefore, in order to seize the opportunity in the fiercely competitive Internet financial market, major platforms often invest heavily in various lucrative activities to attract users. Summarizing the previous literature discussion and empirical research, we propose specific targeted anti-fraud prevention and control countermeasures for Internet financial platform users from the three directions of improving the level of Internet financial supervision, using financial technology to combat fraud, and building industry moral order [8]. Internet platforms provide prevention and control suggestions to reduce company losses and provide more ideas and methods for Internet financial platforms to combat fraud [9].

2. Related Works

According to the literature, the identification of mobile phone users based on biometrics as well as the identification of the use of passwords and patterns is not appropriate [10]. Therefore, the focus of the research is to be able to simply access and recognize the mobile phone sensor data. In order to solve the problem of low stability of dynamic matching recognition, the literature proposes a combination of the D-S evidence theory and mobile phone 3D gesture sensor data [11]. Related researchers have proposed a method called OpenSame to collect data from 200 experimenters in a total of 389,373 tuples [12]. In order to perform implicit authentication of mobile phone users in a sustainable manner without interrupting the normal use of mobile phones, Lee et al. proposed a system based on multiple sensors [13]. Experimental results show that this method is more efficient, the training time is less than 10 seconds, and the verification time is less than 20 seconds. The literature proposes that a platform that provides loans to applicants through the Internet is called an online lending platform [14]. The literature proposes that traditional financial services rely on Internet technology to operate, which can be understood as "finance + Internet".

3. Research and Implementation of User Identity Recognition System in Wireless Sensor Network

3.1. Biometric Technology for Wireless Sensor Network Users. With the rapid development of information globalization in today's society, electronic equipment has become an indispensable part of scientific research and daily life. These devices provide many convenient functions through network interconnection, but the digitization and invisibility of user ID caused by network information is a serious privacy leakage problem [15]. As a result, the requirements for ID recognition

and verification technology have become higher [16]. Among them, the more commonly used methods are security authentication methods including user name and password combined authentication, user-specific object authentication, PIN code authentication, and so on. These authentication methods are easy to be stolen or imitated due to excessive human participation factors and cannot achieve a high degree of security prevention and control. It is very difficult to avoid the dangers brought about by these objective factors [17]. At the same time, some studies have shown that the user's biometric function is difficult to change, imitation and tampering are not easy, fingerprint recognition, facial recognition, iris recognition, movement feature recognition, etc., are very unlikely to be stolen. Figure 1 shows the application process of high-end biometric technology that combines voice and skull fingerprints.

Traditional identification methods usually use identification items (striped codes, QR codes, magnetic cards, keys, etc.) and identification knowledge (passwords, etc.). Because the detection is dependent on foreign objects, if the identification items and identification knowledge are stolen or forgotten, these identities can be easily deceived or replaced by others. Because the biological characteristics are not easy to forget, the anticounterfeiting performance is good, and it is not easy to be tampered with and theft. It can be used at any time by taking advantage of the "portability." The characteristics of the human body are inherently unique to biology. Compared with traditional methods, biometric technology has improved security, privacy, and convenience because keys cannot be copied, stolen, or forgotten. At present, products related to biometric technology are realized through the latest computing technology, combined with the security surveillance system, to realize automatic ID identification and personal information management. According to the current statistical results of the International Biometric Group (IBG), the existing biometric technology is mainly divided into two applications: physiological characteristics and action characteristics, most of which are identification applications that use the characteristics of human physiology. For example, fingerprint recognition, facial recognition, and iris recognition as well as biometric technology based on user behavior characteristics, have multiple applications such as handwriting recognition, voice recognition, gesture recognition, and walking recognition. The indexes of these technologies are shown in Table 1. The biometric method based on action characteristics has the highest safety level and accuracy at the same equipment cost level, high convenience and stability, and low machine cost requirements. Therefore, it has become a new trend in the research field to use the user's behavior characteristics to authenticate the ID of the biometric technology and solve the traditional identification method in the process of vulnerability authentication.

3.2. Visual Sensor Network User Identification Algorithm

3.2.1. Self-Adaptive Recognition Algorithm for Identity Features. The target area environment is defined as a rectangle, and the grid lines are divided into m rows and n
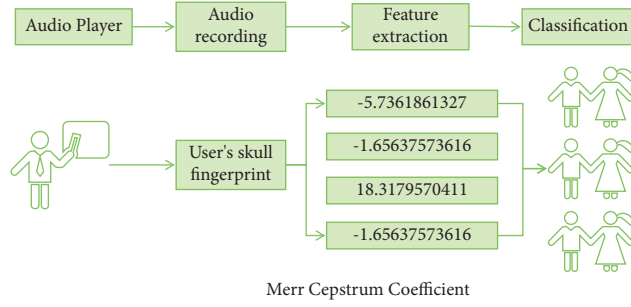
Figure 1: High-end biometric technology using a combination of voice and skull fingerprints.

Table 1: Comparison of various biometric technology indicators.

Biometric technology	Convenience	Accuracy	Security level	Stability	Identify equipment cost
Fingerprint recognition	Higher	High	Medium	Higher	Medium
Face recognition	Extremely high	High	High	Higher	Medium
Speech recognition	High	Medium	Higher	Medium	Lower
Behavior feature recognition	Higher	Higher	Higher	Higher	Medium

columns. According to the corresponding matrix area detection requirements, the corresponding area detection matrix refers to the expansion matrix of all sensors. Define the corresponding restriction matrix suitable for the deployment position of the sensor. By analyzing the number and depth of the recognizable features of the face, the correlation level between the faces can be determined. In this paper, the pixel statistics iterative method is used, and the expression (1) is used to obtain the recognizable average offset vector of the face.

$$S_k(x) = \frac{\sum_{i=1}^{q} E(x_i - x/k)s(x_i)}{\sum_{i=1}^{q} E(x_i - x/k)s(x_i)} - x. \tag{1}$$

After the average offset feature vector of the face is extracted, in order to further improve the effect of identity recognition in the visual sensor network, it is necessary to perform gray-scale processing on the obtained image and apply it to the adaptive recognition of identity features. The specific steps are as follows:

$$\begin{aligned} g(x, y) &= T[f(x, y)] \\ \text{or } D' &= T(D). \end{aligned} \tag{2}$$

Assuming that the gray histogram distribution of the recognized original image is described by $P(x)$, then (2) can

be substituted into $P(x)$ and the gray of the original image can be converted using

$$y = g(x, y) = \frac{1}{c} \int_{x_0}^{x} P(x) dx . \tag{3}$$

The original image after gray-scale conversion was assumed to has a gray level of L and a resolution of $M \times N$. The gray level constant c calculated is as follows:.

$$c = \frac{M \times N}{L - 1}. \tag{4}$$

By substituting (4) into (3) and using (5) to obtain the histogram equalization function, the gray-scale distribution of the image can be made uniform, and the histogram equalization process can be realized.

$$y = \text{INT}\left[\frac{1}{c} \sum_{x_0}^{x} P(x)\right] = \text{INT}\left[\frac{(L-1)\sum_{x_0}^{x} P(x)}{M \times N}\right]. \tag{5}$$

First, a cost function for each gray-level pair corresponding to two images was defined, and the degree of consistency between the gray level at the position y and the gray level at the same gray level was described. Formula (6) is used to explain that the function is expressed as

$$C_p(y) = (1 - \alpha) \cdot \min(d_{iut}(y), \tau_1) + \alpha \cdot \min(d_{Fan}(y), \tau_2),$$

$$d_{\text{int}}(y) = \xi \cdot \frac{\sum_{\delta \in \Omega}\left[(I(y + \delta) - \overline{I}(y)) - (I'(y + \delta) - \overline{I'}(y))\right]^2 \times C_p(y)}{\sum_{\delta \in \Omega}(I(y + \delta) - \overline{I}(y))^2 + \sum_{\delta \in \Omega}(I'(y + \delta) - \overline{I'}(y))^2}, \tag{6}$$

$$d_{sra}(y) = \nabla I'(y) - \nabla I(y) \times C_p(y).$$

In this paper, $L(x, y)$ is used to represent the distance between two nodes (pixels) on the minimum gray tree (x and y), that is, the sum of the weights of the edges connected to the pixels. When the distance between two nodes in the minimum gray level difference tree is very close, the similarity of the corresponding two pixels will increase. On the contrary, if the distance between two nodes of the minimum gray level tree is very far, the similarity of the corresponding two pixels is very small. According to the abovementioned principle, formula (7) is used to calculate the similarity between two pixels.

$$S(x, y) = S(y, x) = \exp(-L(x, y)). \tag{7}$$

Regarding the structure of the minimum gray difference tree, the cost function previously defined for each node (pixel) is expanded to a cost function that converges to the entire tree $T(x)$ represented by

$$C_T(x) = \sum_{v \in F(x)} S(x, v) C_q(v). \tag{8}$$

Similarly, this article uses $C_{ST}(x)$ to describe the total convergence cost in the subtree $ST(x)$ with pixel x as the root node. At this time, the parent node $Q(x)$ of the node is as follows:

$$
\begin{aligned}
C_{ST}(x) &= \sum_{v \in ST(x)} S(x, v) C_q(v) \\
&= C_q(x) + \sum_{v' \in ST(x)} S(x, v') C_q(v'') \\
&= C_q(x) + \sum_{v' \in ST(x)} S(x, v') S(v', v'') C_q(v'') \\
&= C_q(x) + \sum_{Q(v')=x} S(x, v') \sum_{v' \in ST(v')} S(v', v'') C_q(v'') \\
&= C_q(x) + \sum_{Q(v')=x} S(x, v') C_{ST}(v').
\end{aligned}
\tag{9}
$$

The parent node of the subtree is used to perform line-by-line design for each layer, and formula (10) is used to establish the minimum gray-scale tree model.

$$C_{ST}(x) = C_q(x) + \sum_{Q(v')=x} S(x, v') C_{ST}(v'). \tag{10}$$

According to the matching cost calculation, the matching cost of the entire recognized image (size H•V) is obtained by adding up and normalizing at the same time, and finally, the matching cost (MC) of the image and the similarity of the two images are obtained.

$$MC(I, I') = \sum_h \sum_w C_T \frac{(x)}{(h \cdot w)}. \tag{11}$$

3.2.2. Support Vector Machine. The basic idea of a support vector machine (SVM) is to divide the data correctly. Therefore, important parameters such as the penalty

coefficient C and the kernel width γ of the kernel function must be set to appropriate values before applying actual problems. The penalty factor C plays an important role in minimizing fitting errors and balancing the complexity of the model. In the traditional method, these parameters are processed according to the grid search and the fastest descent method. However, these methods can be simply classified as local optimal solutions. In recent years, several meta-heuristic search algorithms based on Biological Hughstick have appeared. It is easier to find the best global solution than the traditional methods mentioned above. Based on the above analysis, an improved gray wolf optimization algorithm is proposed to optimize the selection of SVM model parameters. This method is suitable for the problem of crop lack of diagnosis.

In the sample space, the hyperplane of the sample can be divided by a linear equation.

$$w^T x + b = 0. \tag{12}$$

The support vector is constrained to be correct. That is, the support vector satisfies the following conditions.

$$y_i(w^T x_i + b) - 1 = 0. \tag{13}$$

Corresponding to the function interval, the geometric interval has better properties. The geometric interval of the sample points is defined as

$$\gamma_i = y_i \left(\frac{w^T}{\|w\|} x_i + \frac{b}{\|w\|} \right). \tag{14}$$

The geometric interval of the entire training sample set is defined as the minimum value of the geometric interval of all sample points.

$$\gamma = \min_{i=1,\dots,N} \gamma_i. \tag{15}$$

This is regarded as the original optimization problem, the Lagrangian pairing is applied, and the Lagrangian multiplier is added under the constraints of the equation. Therefore, the Lagrangian function of this problem can be expressed as follows:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i \left(y_i (w^T x_i + b) - 1 \right). \tag{16}$$

When all constraints are met, there are

$$\theta(w) = \frac{1}{2} \|w\|^2. \tag{17}$$

In other words, Equation (17) constitutes the basic form of SVM. Therefore, if satisfying all constraints, $\theta(w)$ is the same as direct minimization, and the objective function can be expressed as follows:

$$\min_{w,b} \theta(w) = \min_{w,b} \max_{\alpha_i \geq 0} L(w, b, \alpha) = p* \tag{18}$$

Swap the positions of the minimum and maximum functions to obtain the following information:

$$\max_{\alpha_i \geq 0} \min_{w,b} L(w, b, \alpha) = d*. \tag{19}$$

The conditions of the Lagrangian multiplier method was satisfied, and the partial derivative function of each component was taken, when the value is 0:

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i,$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0. \tag{20}$$

As shown in (21), the parameters w and b were deleted, and only the problem was converted into a variable.

$$\max_{w,b,\alpha} L(w, b, \alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j. \tag{21}$$

3.3. Wireless Sensor Network Data Collection and Processing

3.3.1. Collection of Geometric Features of Human Faces.
The core function of the system is the display of sensor information, and it is necessary to design a table for initial storage of data. SQLSERVER was choosen as a stable and reliable high-performance database management system because it has a large-capacity data storage system, a variety of data, and long-term storage time period characteristics. The design of the data table must be reasonable and comprehensive, the image of each person in the neutral facial image was registered to the visual sensor network, and the other three representation variables and three illumination variables were used as the detection image set. The collected functions are shown in Table 2.

After obtaining the correct geometric features of the face, it is necessary to collect the geometric features of various facial expressions. The detailed expression conversion is shown in Table 3.

3.3.2. Some Features Used in Gait Recognition.
Mark Nixon, a professor of electrical engineering and computer science at the University of Southampton in the United Kingdom, found that the way people walk is fundamentally different. There are subtle aspects in muscle strength, flexibility of tendons, length and density of bones, vision, and damage to muscles and bones. Difference: the walking process of a normal adult is based on the regular movement of taking the left foot one step forward and the right foot one step forward, which is called the complete walking cycle. In this chapter, it is divided into several short periods according to the adjacent minimum value on the Y axis of the acceleration sensor, so that it can easily determine the walking period corresponding to the data. According to statistics, the average time of a walking cycle is 1 to 1.5 seconds. These are based on complete walking time. The main characteristics considered are shown in Table 4.

TABLE 2: Geometric features of standard human faces.

Feature	Geometric characteristics	Measurements
F1	Inside and outside eye point width/cm	3.9
F2	Distance between inner eye points/cm	4.2
F3	Outer eye point interval/cm	10.3
F4	Height of nose/cm	1.4
F5	Nose tip angle of left and right wings	116°
F6	Eye point and nose angle in left and right	32°
F7	Nose volume/cm^3	7.9

TABLE 3: Acquisition of geometric characteristics of the 3D face model.

Feature	Smile	Frustrated	Pouting	Bulging
△F1/cm	0.9	1.5	0.2	0.2
△F2/cm	0.5	1.8	0.3	0.3
△F3/cm	3.5	2.6	0.4	0.3
△F4/cm	0.6	0.3	0	0
△F5	6°	5°	0	0
△F6	4°	10°	2°	4°
△F7	25°	8°	6°	5°
△F8	45°	26°	25°	25°
△F9/cm^3	4	2.6	0	0
Geometric similarity	0.58	0.45	0.60	0.55
Relevance similarity	0.64	0.71	0.63	0.59

3.3.3. Synchronization.
In theory, when the sampling frequency of different sensors is set to the same, the data needs to be collected at the same time. In other words, the time stamps of the sampling points are exactly the same. However, in actual situations, even if the same smartphone is set to the same sampling frequency, the returned sensor data will have a different sampling time. Therefore, in order to avoid the time difference between the same frame of data intercepted according to the length of different sensors, it is necessary to synchronize the original sensor data. This process is shown in Figure 2.

3.4. Simulation Experiment Results and Analysis

3.4.1. Evaluation Index.
According to the definition of the object problem, the user identification of the smartphone belongs to multiple classification problems. The model classification performance evaluation index proposed in this paper uses accuracy, F1 value, ROC curve, and AUC. The calculation formulas for the matching rate and reproduction rate of the k category are as follows:

$$P_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FP}_k},$$

$$R_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FN}_k}. \tag{22}$$

Accuracy is the most common evaluation index for classifiers. This is the ratio of the number of samples that are correctly identified (that is, correctly classified) for all sample sizes. For example, there are 1,000 students in junior high

TABLE 4: Some features used in gait recognition.

Feature	Description
Root mean square	The sum of squares in the three directions of X, Y, and Z is divided by 3, and the square is taken
Energy	Work done in the three directions of X, Y, and Z
Percentile value	The percentage of the number of samples below this sample value in the dataset to the total number of samples
Mean absolute deviation	The average of the absolute value of the deviation of each value from its arithmetic mean
Cadence	The number of complete gait cycles contained in 1 minute
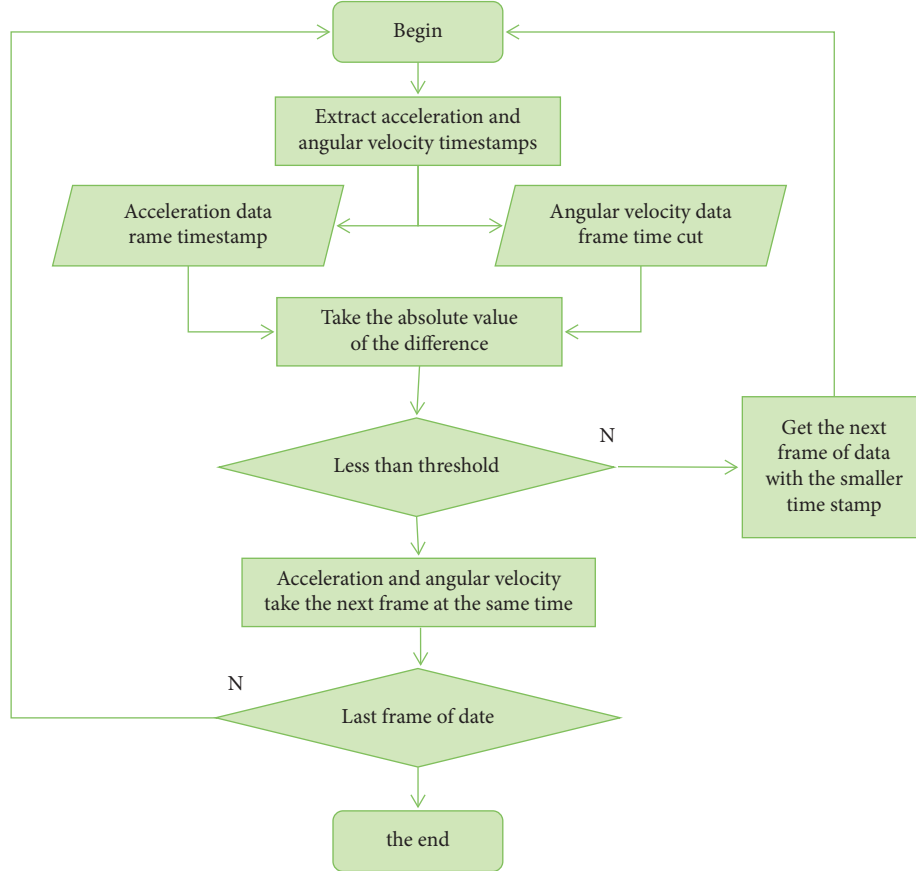
FIGURE 2: Flow chart of sensor data synchronization process.

school. Obviously, there are two categories of students, boys and girls. Among them, 600 are boys and 400 are girls. If the classifier correctly determines the gender of 800 people (including 500 boys and 300 girls), then, the number of people correctly classified by the classifier will account for 80% of the total. In other words, the accuracy is 80% (800/1000). This paper shows the calculation method in following formula:

$$\text{accuracy} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{sampls}-1} f(\widehat{y}_i = y_i). \qquad (23)$$

The F value is the weighted harmonic average of precision and precision. For multiple classification problems, the evaluation indicators of micro-F score and macro-F score are usually used. Because the coincidence rate and the reproduction rate may be mutually exclusive, it is necessary to consider these two issues together. In order to solve this

problem, the F value is proposed, and the calculation formula is as follows:

$$F_\beta = \frac{\left(1 + \beta^2\right) \times P \times R}{\beta^2 \times P + R}. \qquad (24)$$

3.4.2. Detailed Comparison of Recognition Performance. In order to further prove the effectiveness of the algorithm in this paper, on the basis of the abovementioned experiment, the performance of the two algorithms is statistically compared, and the results obtained are shown in Table 5.

3.4.3. Division Ratio Results. Figure 3 shows the experimental results of the ShakeLogin dataset under experimental conditions. For the model SSUI proposed in this paper, when the division ratio is 0.8, the accuracy rate reaches the

TABLE 5: Detailed comparison of the performance of the two algorithms.

Variable type	The recognition accuracy of the algorithm in this paper (%)	Recognition accuracy based on palmprint and facial feature algorithm (%)
Happy	100	84
Sad	97	76
Surprised	96	62
Light source 1	98	76
Light source 2	96	69
Light source 3	90	63
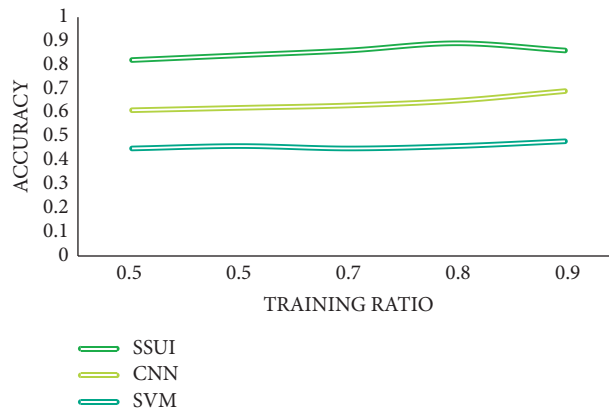
FIGURE 3: Performance graph of the ShakeLogin dataset with different division ratios.

maximum, and then, when the division ratio is increased to 0.9, the accuracy rate drops slightly.

The specific experimental results are shown in Figure 4. It can be seen from the broken line chart that as the proportion increases, the accuracy of walking and going up and down stairs increases, and the accuracy of going up and down stairs increases by more than 20%. The other three actions have changed, but when the ratio reaches 0.9, and all actions have reached the maximum accuracy.

3.4.4. Results of All Behavioral Experiments. The results of this series of experiments are shown in Table 6. The time-domain feature weighting parameter in this table changes from 1 to 2/3, 1/2, 1/3, and finally to 0. The frequency domain weighting parameter value ranges from 0 to 1/3, 1/2, 2/3, and finally to 1.

Figure 5 shows the performance of the SSI model in the Shakelogin dataset and the experimental results of the sample period. As shown in the figure, the sample period is too short (2 seconds), the accuracy is very low, and the obtainable value is less than 60%. The longer the sample time is, the higher the correct rate (2 seconds to 4 seconds, 60% to close to 80%) is.

4. User Identification Applications and Risk Prevention and Control Strategies in Internet Financial Platforms

4.1. Identification of Fraudulent Users in Internet Financial Platforms. The data in this article comes from the user transaction record data of a European online payment platform, which was collected and shared during the research cooperation process of the machine learning team of Worldline and ULB (University Libre de Bruxelles). Because each transaction data contains a category label (that is, whether it is a fraudulent transaction), this is actually a supervised classification machine learning task. This experiment will first perform data preprocessing, that is, subsamples were obtained by random undersampling, and then, feature engineering was performed, including feature scaling, outlier detection, and feature screening. Subsequently, four supervised machine learning algorithms are used to train and test the model, and finally, a variety of evaluation indicators are used to evaluate the effects of various models.

4.2. The Main Links of User Identification in Internet Financial Platforms

4.2.1. Verifying Customers. Confirm the customer and its actual manager, understand the natural purpose and transaction, confirm the identity of the customer, understand the actual beneficiary of the natural person and the actual management of the transaction, confirm the customer's valid ID card or other ID, register the customer's basic identity information and Relevant ID card copy information. This basic task is particularly important. The main reason is that in order to avoid legal sanctions when performing money laundering activities, criminals must
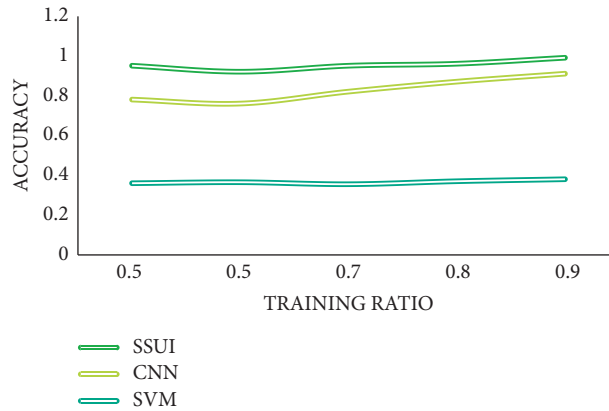
FIGURE 4: HHAR dataset (all behaviors).

TABLE 6: User identification results of ShakeLogin and HHAR datasets.

ShakeLogin (17 users, shake the phone arbitrarily)	Random 80% of the entire data set/remaining	85.46	87.25	92.07	81.98	75.33
HHAR (9 users, 6 behaviors)	All actions/all actions	92.56	97.12	92.56	97.23	94.88
	All behaviors/walking	90.08	96.87	97.75	96.85	100.0
	All actions/stations	90.37	97.78	95.84	98.05	72.42
	All behaviors/sit	86.18	90.55	92.67	93.83	66.05
	All behaviors/biking	90.92	85.32	91.62	88.83	86.37
	All actions/stairs	92.75	96.93	91.35	95.16	97.59
	All actions/down stairs	88.73	93.18	95.59	94.39	97.56
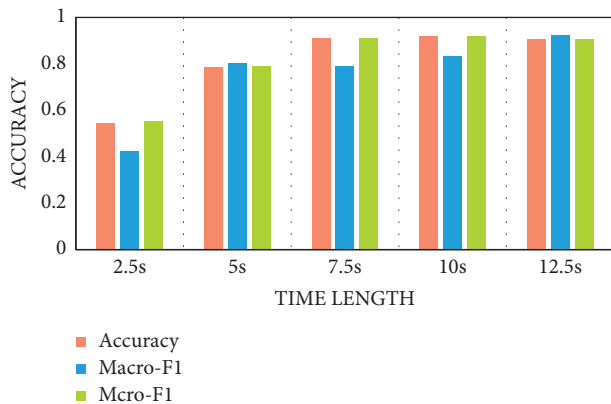
FIGURE 5: SSUI's performance in the ShakeLogin dataset varies with sample duration.

conceal their true identity and actual transaction purpose. In order to effectively identify suspicious personnel and suspicious transactions of financial institutions, it is necessary to strictly and carefully determine and verify the true identity of customers and pay attention to daily business and financial transactions to ensure that abnormal situations are discovered in a timely manner.

4.2.2. Continuous and In-Depth Identification of Customers. During the duration of their business, financial institutions also need to continue to reexamine customer accounts and transaction activities. This process is actually a process for financial institutions to learn more about customers and confirm their identities. In order to effectively achieve continuous identification, financial institutions need to do the following.

As a work to prevent money laundering, the most basic thing is to be able to update customer information and related data in a timely manner. The staff at the counter of the basic relevant outlets needs to have sufficient sensitivity and attention to customer information and customer managers through early learning of the contact process of customer information. Then, continue to pay attention to the latest trends of customers, understand and update the money in the customer's account, and the actual owner or beneficiary of these funds. In the absence of a legal person or other legal person organization, the credibility of relevant information must be judged carefully. In order to understand and prevent all the customer's structure and internal management structure from being false, it is necessary to use the help of the propaganda platform to check and verify the offender as much as possible, to ensure that customers do not use false identities to hold stock or fund accounts and engage in money laundering activities.

Please pay attention to the customer's transaction behavior. Please note that whether the money in the customer's account flows regularly or is very different from other accounts of the same or similar type. The average daily occurrence of the account, the degree of frequent trading operations, the source of remittance funds and account outflows, and the fixed transaction objects. Please pay attention to capital accounts with common characteristics. Suspicious accounts such as frequent and rapid fund transfers, transactions of almost equal amounts between borrowers and credit cards in a short period of time, and unusually active account transactions.

Conduct customer surveys on a regular basis. In this work, the relevant staff must be familiar with the relevant business information of the customer, understand the industry in which the customer is engaged in advance, master the general financial settlement methods in the business process and whether they prefer cash settlement. In the practice of preventing money laundering, criminals frequently use cash transactions in order to avoid transfer of funds between banks or use promissory notes in the same city according to the needs of illegal and criminal activities. The information of the recipient, payer, and correspondent bank cannot be clearly displayed on the same bill, hide the source and location of funds, or personal accounts or frequent electronic remittances between units and personal accounts. Therefore, financial institutions cannot understand the actual parties behind the transaction and make decisions in a timely manner. Finally, it is necessary to strengthen system construction, establish customer electronic information database, establish customer credit database, and integrate system platform resources.

4.2.3. Due Diligence on High-Risk Customers. Usually, the age and occupation of the customer obviously do not match the transaction amount and transaction frequency or customer accounts belong to high-risk industries, such as offshore companies or entertainment facilities, or individual customer accounts registered on the blacklist, such as terrorists or politically sensitive. Digital and financial institutions must conduct detailed due diligence on them. The customer's due diligence must be carried out through the whole process of the financial institution's business start, decision-making, and postloan management. The content of the investigation includes the complete collection of the client's ID card, the understanding of the actual account manager or the actual beneficiary of the transaction, the confirmation of the client's valid ID or other IDs, the understanding of the source of funds, the use of funds, economic status or business status, and other information.

4.3. Development Strategies for Risk Prevention and Control of Internet Financial Platforms

4.3.1. Start with the Details of the Process, and Comb Out the Operating Guidelines and Implementation Standards Based on the Characteristics of Your Own Business. Financial institutions need to incorporate customer identification into the internal control construction system, improve operating standards, clarify business processes, and implement customer risk classification and related risk classification standards. Strengthen the specific conditions and recognition points of initial recognition, continuous recognition, and rerecognition. Government regulatory agencies, the People's Bank of China, financial institutions, public security, taxation, legal affairs, commerce, and foreign exchange administration departments provide advice and relevant professional support for the effective identification of various client ID information of financial institutions. Build a unified comprehensive information query system to solve the information problem of financial institutions' verification practices.

4.3.2. Make Full Use of Information Technology and Strengthen Identification Methods. One is to maximize the use of existing professional databases, such as World Check, to screen potential risks for customers, partners, transactions, and employees. World Check is public in the world. It contains public data on money launderers, scammers, terrorists, drug smugglers, fraud companies, etc. It is a recognized organization database. The second is to strengthen the use of advanced technologies such as fingerprint recognition, iris detection, and keyboard rhythm. At the same time, in order to improve the recognition accuracy of Internet financial services, in addition to general passwords and digital certificates, it is also necessary to increase the SMS and voice recognition of mobile phones. The third is to integrate the input and integration of scattered customer ID information, combine industries, regions, locations, other dynamic IDs, and automatic screening of high-risk customers to develop ID information tracking, monitoring, and analysis systems. With the rapid development of Internet technology, the conflict between traditional financial commerce and emerging Internet financial commerce has brought more difficulties and problems to the prevention of money laundering. In the new situation, financial institutions must respond in a timely manner. In order to maximize the use of the advantages of Internet financial services, effective countermeasures are taken on the basis of ensuring the identification to the greatest extent to achieve a balance between risks and benefits.

5. Conclusion

Based on wireless sensor networks, this paper studies the development of user identification and network finance. This paper studies the characteristics of wireless sensor data, better data reduction, and data classification algorithms and proposes a basic action recognition method based on wireless sensor data. Aiming to improve the shortcomings of traditional identification algorithms, an adaptive identification algorithm based on the improved minimum gray tree of the visual sensor network is proposed. The simulation results show that the algorithm shows a high accuracy of identification. On the other hand, wireless sensor call is simple and easy to obtain data, so it forms the basis of wireless sensor data research. Each user who uses the wireless sensor generates sensor data, and the sensor data is generated in different ways depending on the user. Therefore, the user can be identified through wireless sensor data. The prevention of money laundering is affected by the rapid development of Internet technology and the conflict between traditional financial commerce and emerging Internet financial commerce, and many new problems have emerged. Financial institutions in a new situation need to respond in a timely manner. The concept of the effectiveness and maximum guarantee of the countermeasures is to maximize the use of the advantages of Internet financial services to achieve

a balance between risks and profits. The establishment of a complete and effective customer information database will undoubtedly increase the human and material resources of various financial institutions. However, there are many problems such as information collection and information verification, and huge investment may be effective, which will have a certain degree of impact on the optimization and improvement of "customer identification."

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[2] X. Huang, Y. Xiang, A. Chonka, J. Zhou, and R. H. Deng, "A generic framework for three-factor authentication: preserving security and privacy in distributed systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 8, pp. 1390–1397, 2010.

[3] Y.-H. Jo, S.-Y. Jeon, J.-H. Im, and M.-K. Lee, "Vulnerability analysis on smartphone fingerprint templates," in *Lecture Notes in Electrical Engineering*, vol. 354, pp. 71–77, Springer, Berlin, Germany, 2016.

[4] C. Wang and G. Xu, "Cryptanalysis of three password-based remote user authentication schemes with non-tamper-resistant smart card," *Security and Communication Networks*, vol. 201714 pages, 2017.

[5] F. Wang, G. Xu, and G. Lize, "A secure and efficient ECC-based anonymous authentication protocol," *Security and Communication Networks*, vol. 201913 pages, 2019.

[6] K. Wang, S. Tsai, X. Du, and D. Bi, "Internet finance, green finance, and sustainability," *Sustainability*, vol. 11, no. 14, p. 3856, 2019.

[7] L. I. Li-wei and J. Feng, "Relationship between internet diffusion and economic growth: empirical research based on panel data of China's 31 provinces," *Journal of Beijing Technology and Business University*, vol. 28, pp. 120–126, 2013.

[8] D. Georgarakos and G. Pasini, "Trust, sociability, and stock market participation," *Review of Finance*, vol. 15, no. 4, pp. 693–725, 2011.

[9] Y. Sahin, S. Bulkan, and E. Duman, "A cost-sensitive decision tree approach for fraud detection," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5916–5923, 2013.

[10] A. K. Das, "Analysis and improvement on an efficient biometric-based remote user authentication scheme using smart cards," *IET Information Security*, vol. 5, no. 3, pp. 541–552, 2011.

[11] C. C. Chang, S. C. Chang, and Y. W. Lai, "An improved biometrics-based user authentication scheme without concurrency system," *International Journal of Intelligent Information Processing*, vol. 1, no. 1, pp. 41–49, 2010.

[12] Y. An, "Security analysis and enhancements of an effective biometric-based remote user authentication scheme using smart cards," *Journal of Biomedicine and Biotechnology*, vol. 2012, Article ID 519723, 6 pages, 2012.

[13] J. K. Lee, S. R. Ryu, and K. Y. Yoo, "Fingerprint-based remote user authentication scheme using smart cards," *Electronics Letters*, vol. 38, no. 12, pp. 554-555, 2002.

[14] Y. Liu and H. Wang, "The risk evaluation and control of private lending in China from the perspective of Internet: taking P2P platform as an example," *Macroeconomics*, vol. 3, pp. 146–157, 2017.

[15] Y. Y. Cheng, L.-Y. B. Ying, S.-B. R. Jiao, P.-R. G. Su, and D.-G. Feng, "Research on user privacy leakage in mobile social messaging applications," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 37, no. 1, pp. 87–100, 2014.

[16] Z. F. Huo, X.-F. Meng, and Y. Huang, "PrivateCheckIn: trajectory privacy-preserving for check-in services in MSNS," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 36, no. 4, pp. 716–726, 2013.

[17] S. Park, J. Byun, J. Lee, J. H. Cheon, and J. Lee, "HE-friendly algorithm for privacy-preserving SVM training," *IEEE Access*, vol. 8, pp. 57414–57425, 2020.

Research Article

Intelligent Prediction and Rural Financial Development Based on Abnormal Detection of Sensor Data

Jiong Liu [ID]

Department of Tourism and Commerce, Xuancheng Vocational and Technical College, Xuancheng, Anhui 242000, China

Correspondence should be addressed to Jiong Liu; liuj@xcvtc.edu.cn

Wireless sensor network is a multisensor wireless network system, which consists of multiple sensors and is configured independently. Because the network generates a large amount of data, the frequency, performance, and computing power of sensor nodes are limited, and they are particularly vulnerable to harsh environments and malicious attackers. This leads to the occurrence of malicious nodes, emergencies, and abnormal data in the sensor network system. Failures can also have a significant impact on sensor network services. The two main functions of wireless sensor network security are abnormal node detection and data anomaly detection. These two directions are mutually independent and complementary. Therefore, under the promotion of the rural revitalization strategy and the precision poverty alleviation strategy, China has increased its agricultural efforts. At this stage, all localities focus on the construction of rural financial systems to ensure that scattered farmers and rural small and microenterprises receive comprehensive financial services. The establishment of a rural financial system based on "intelligent forecasting" can improve financial development theories and build new ideas for rural financial development. And, the balance between realization and profitability, and then, through the use of Internet technology to make traditional financial institutions more effective in providing financial services, new online financial platforms can use them to make up for the existing shortcomings of traditional financial institutions as much as possible. In this article, through the research on the intelligent prediction of sensor data anomaly detection, it is applied to the development of rural finance and promotes the development of rural finance.

1. Introduction

In recent years, with the rapid development of computer technology, electronic equipment, wireless networks, and communication systems, wireless sensor networks have continued to emerge and have gradually become an important channel for people to obtain data and information [1]. Wireless sensor networks have the characteristics of limited node sources and are easy to break [2]. Although technologies such as mandatory passwords and secure routing have improved the security of sensor networks, there is still a lack of effective methods to detect abnormal information in the sensor network, making the sensor network more effective. Protection becomes difficult [3]. The improvement of data anomaly detection technology can greatly promote the future application and development of sensor

networks. Each sensor node maintains the trust value of neighboring nodes to reflect its previous decision-making behavior [4]. Two thresholds are used to reduce the false alarm rate and increase the incident. Area detection accuracy is used to achieve more accurate detection of malicious nodes without sacrificing normal nodes [5, 6]. The simulation results show that the two-threshold scheme is better than the single-threshold scheme [7]. After screening and removing malicious nodes in the sensor network, a fault detection method based on a distributed multilayer wireless sensor network is introduced [8]. Through the analysis of false datasets and real-time environmental data collection data, it is shown that compared with the central system and the reference system, the improved anomaly detection program can achieve a higher detection rate, false alarm rate, and lower communication utilization [9]. Finance

conducted a "Internet + Intelligent Forecasting" research. Using the concepts and theories related to the overview, firstly, it analyzes the status quo of rural financial development under the background of "Internet + intelligent forecasting" from the service functions of traditional rural financial institutions and the service functions of Internet rural financial institutions; Alibaba and CreditEase analyze and demonstrate the three typical rural finance "Internet + smart forecasting" cases to obtain development information: clarify the "Internet + smart forecasting" agricultural development strategic goals, make full use of big data, and promote the self-improvement of local financial institutions; By further analyzing the problems existing in the development of rural finance in the context of the "Internet," we found that there are various problems in rural financial risks. It is necessary to further expand the rural credit control system, improve Internet finance laws and regulations and rural network infrastructure, and increase the penetration rate of farmers' financial knowledge [10]. Finally, it puts forward targeted countermeasures to promote the development of rural finance under the background of "Internet + intelligent forecasting," establishes a political incentive mechanism for state intervention, strengthens the prevention and monitoring of rural financial risks, and promotes financial publicity and education in rural areas [11, 12].

2. Related Works

The literature introduces the application background of wireless sensor networks, describes the research background and practical significance of malicious nodes and abnormal data detection, analyzes the current status of the investigation of malicious messages and abnormal information detection technologies at home and abroad, and provides a complete, structured, and summarized summary [13]. The literature introduced the basic problems of wireless sensor networks, introduced the anomaly detection datasets commonly used in wireless sensor networks, and explained the architecture of wireless sensor networks in detail. The literature introduces a two-step detection system for malicious nodes on wireless sensor networks. Each sensor node maintains the trust of neighboring nodes to reflect its past behavior in the decision-making process [14]. The simulation results show that the two-tier plan is better than the single-tier project. The program provides a set of parameters whose values can be adjusted according to the application system to maximize efficiency. The literature introduces an abnormal data detection method for distributed wireless sensor networks based on hierarchical aggregation [15]. Through simulation analysis on artificially constructed datasets and datasets collected in real environments, it shows improved anomaly detection schemes and centralized schemes. Compared with the HSCBD scheme as a control, it can achieve a high detection rate and a low false alarm rate and reduce communication consumption. The literature introduces a distributed error detection system [16]. The algorithm includes two main steps: self-detection and intermediate pod detection. The algorithm makes full use of

the spatiotemporal data connection of data distribution in wireless sensor networks and divides abnormal data into two steps: time and location [17]. Carry out anomaly detection process on the basis of reliability improvement.

3. Design and Implementation of the Intelligent Prediction Method Based on Abnormal Detection of Sensor Data

3.1. Analysis of Wireless Sensor Network Architecture and System Requirements

3.1.1. Wireless Sensor Network Architecture. A sensor is usually a cheap device with relatively low battery life, computing power, and storage capacity. The main purpose of the network is to identify, store, process, and transmit environmental information in the area covered by the system. Related technologies mainly include three major technologies: sensor technology, information technology, and communication technology. The three architectural components of a wireless sensor network are sensors, sensing objects, and users.

There are many types of wireless sensor nodes, but a single node has a single function. It can generally be used to measure various indicators in the natural environment, such as earthquake, noise, temperature, humidity, and light intensity. It has a wide range of application prospects and is used in various parts of today's society. Figure 1 is a typical wireless sensor network architecture diagram, in which the sensor nerves are automatically distributed in the area of interest manually, and they work together to collect and process the data in the coverage area and transmit all or part of the data to the sewer or base station. After the summary data are sent, the measurement data can be processed into multiple items. The base station sends the metadata collected through satellite or wireless network back to the administrator for decision-making. The administrator can also send instructions to the base station through the network and send the instructions to each relevant sensor.

Most sensor nodes are very small. Because the technology development is not very mature, the computing power, memory, and communication capabilities of the nodes are limited. Network sensor systems are often used in unintentional or difficult environments, and the battery cannot be replaced, and the life cycle is relatively short. Many wireless sensor network power supply modules include wireless communication modules, processing modules, and sensors. With the development of single-chip technology and electronic integration technology, computer power consumption has been greatly reduced. In addition to collecting and processing local data, sensors must also cooperate with other entities to perform certain tasks under special conditions.

The convergence node is an intermediary between the sensor network and the external Internet, and can identify the exchange protocol between the Internet of Things and the Internet, processed by the network sensor, and sent to the management node after collection and processing.
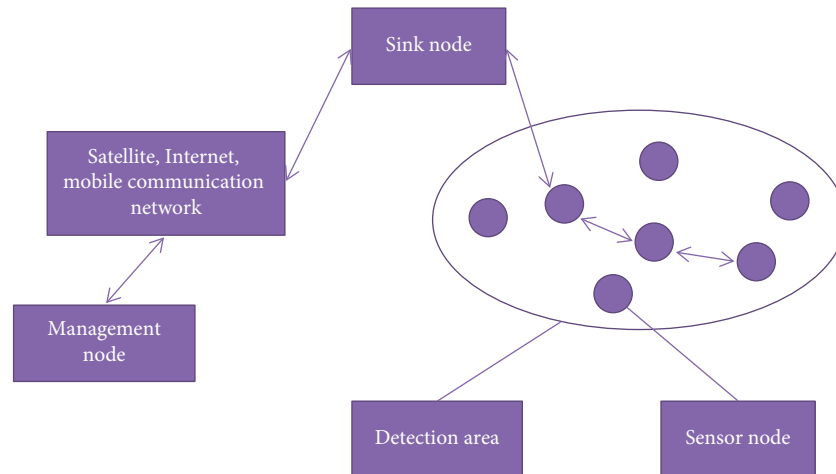
FIGURE 1: Typical wireless sensor network architecture diagram.

The management node is the "stand-in" of the entire sensor network and generally directly controls the sensor identification system of the sensor network through it, and the network administrator can access the review intelligence through management, decision-making, and guidance clauses.

3.1.2. The Internal Node Composition of the Wireless Sensor. The wireless sensor network consists of a large number of randomly distributed sensors. For example, if we conduct an environmental audit, the data will collect various information in the environment (such as temperature, humidity, and air pressure) to form a self-hosted wireless network. Each sensor is sent finally to the management terminal and configuration terminal. There are two types of nodes in the network: ordinary sensor nodes and gateway nodes. Distributed sensor networks can communicate with each other and connect to a gateway node through multiple data. The gateway node collects and transmits data. The network can access the Internet through the gateway node. Communication with the entire site of the data controller and the entire system can be managed and controlled from the headquarters.

The internal structure of the sensor is shown in Figure 2. The data acquisition module is composed of two subunits: the sensor component and the analog-to-digital converter (AD/DC). The digital output signal is transmitted to the data processing module through the analog-to-digital converter. The power supply module supplies power to each sensor module and usually uses a battery to power the device. When designing sensor hardware components, save device power as much as possible and extend the life of the device. It is necessary to use components with the lowest power consumption possible, and the sensor should not communicate. Turn off some communication functions when needed. When developing communication software for sensor networks, you must consider the energy consumption of the network. By eliminating some unimportant network performance indicators, you can achieve higher power utilization and extend the life of the entire network.

3.1.3. Data Security Requirements of Sensor Networks. Wireless sensor networks have different security requirements in different application scenarios. Aiming at sensor network systems commonly used in industrial and agricultural production, this article divides the most important security requirements of wireless sensor networks into basic security requirements and network system security requirements (see Table 1).

The threats to wireless sensor networks are caused by environmental factors such as natural disasters and intruders, and face different threats in different fields such as data, communications, services, and hardware, as shown in Table 2.

Wireless sensor network nodes are large in number and widely distributed. In practice, sensor networks are usually deployed in areas where few or no people are on duty. In addition, the hardware fragility of sensors and some potential software defects cause further damage to the network.

An information attack is an attacker maliciously intercepting, threatening, accessing, or destroying data in the network, resulting in incorrect judgments of network data or errors in data reporting. False information is confusing and difficult to be discovered by people. Therefore, such attacks are very hidden and dangerous and should be considered. In response to such attacks, the sensor network data detection technology cannot analyze the dataset by the wireless sensor network to determine whether the node data are abnormal, thereby reducing the harm of data attacks.

Attackers perform Sybil attacks by invading network nodes and then using authentication technology to impersonate multiple legitimate network nodes. This attack will destroy the network topology and disrupt normal network communication. Vulnerability attacks are mainly related to the attacker manipulating data to be sent after the intrusion part or all of the data are related, resulting in the loss of a large amount of normal data on the network and abnormal communication between messages. Most flood attacks have two attack methods. One attacker controls the node to continue trying to establish a connection with the attacked node to use the attacker's node connection resources, and the other is that the attacker controls the node to
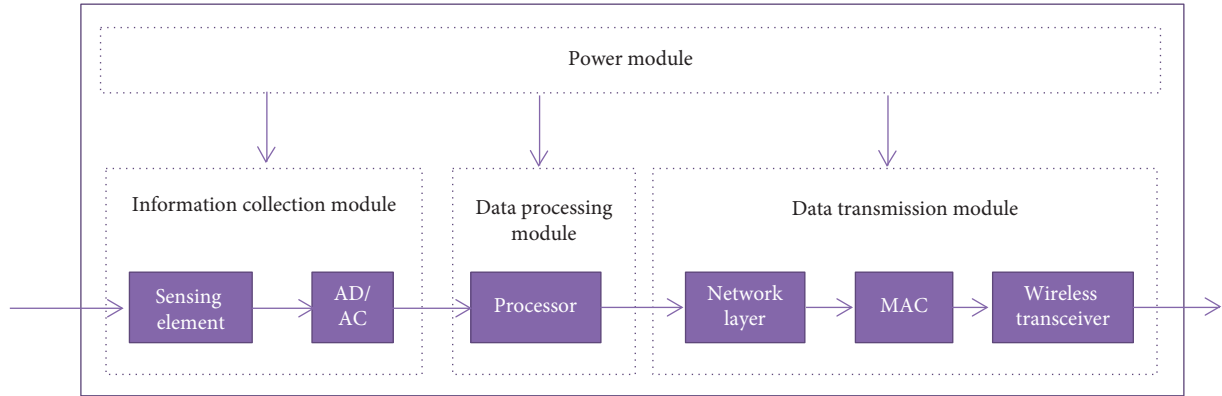
FIGURE 2: Structure diagram of the internal composition of sensor nodes.

TABLE 1: Security requirements of wireless sensor networks.

Species	Feature	Demand details
Basic security requirements this An all	Confidentiality Availability	Ensure that authorized users access information and prohibit unauthorized access. Ensure that legitimate users do not deny the abuse of data and resources.
	Completeness	When inputting and forwarding, the data are not displayed, and it is illegal to agree to modify and destroy the information to protect the consistency of the information and data.
	Authenticity	Ensure the authenticity and credibility of the identities and information of both parties communicating with each other.
Need	Data security Self-healing and recovery ability	Unauthorized access, modification, or transmission of any data is not allowed. The network must have a certain degree of self-defense ability to resist unknown attacks.

TABLE 2: Security threats of sensor networks.

Field	Main threat
Data	Illegal eavesdropping, malicious data tampering, false data injection
Communication	Witch attack, black hole attack, flood attack, wormhole attack
Service	DDOS attack
Hardware	Theft, destruction, equipment hardware intrusion, natural disasters

continuously send invalid data to the attacked node. The request is forwarded by the attacked node, thereby destroying the computer and frequency band node resources. Both of these attack methods prevent the attacker from processing legitimate requests from other data and greatly increase the energy consumption of the attacked node. Most wormhole attacks target two malicious nodes that together form a communication tunnel. Even if the two malicious nodes are far apart, there is only one step between the two malicious nodes. In this way, the hop times of the two nodes are shorter than normal, so as to achieve the purpose of deceiving the right of way and destroying the normal routing path.

Service attacks are mainly carried out through distributed denial of service (DDOS) attacks. The purpose is to exhaust the communication resources or computing resources of the sensor network, causing the system service group network to be unavailable.

The purpose of hardware attacks is to control smart devices in a sensor network by destroying device hardware or intruding devices. Direct damage to such devices will

cause some system functions to be paralyzed. Smart device monitoring can be used as a spring trigger to attack sensor network systems, such as managing sensor data through data transmission, sending false monitoring data to attack network readiness, running malicious code to attack other devices, and preventing hardware attacks by installing protective shells and other security measures. It can prevent the physical security of network equipment. Using fault detection technology, a strict access control system can be established to detect faulty equipment in time and prevent intruders from entering the equipment.

Anomaly attributes are various network status indicators that appear when the sensor network is threatened. In order to use the K method to detect anomalies, information objects need to be grouped according to the anomaly attributes. Therefore, the basic principle and description of character selection is to confirm the K-mean algorithm detection result important factor.

Aiming at the threat of data attack, through analyzing typical data attack methods, the selected characteristics are shown in Table 3.

TABLE 3: Attack characteristics of sensor networks.

Typical attack method	Attack description	Representative characteristics
Malicious data tampering	Attackers maliciously modify the collected data or reported data in the network, causing abnormalities in the data transmitted and reported by the network nodes.	Perceived data such as temperature, humidity, luminosity, and data flow information
Fake data injection	Attackers use false control data or unsupervised data to compress the sensor network, causing network failures or unreliable data reported by nodes.	Sensing data such as temperature, humidity, luminosity, and data flow information entropy

The data collected by sensor nodes generally have several dimensions. Through the analysis of multiformat databases, the spatial interaction of sensor network data can be effectively used to detect data errors. The data can reflect the feasibility and statistical characteristics of sensor data flow. By introducing the work of this group of analysis, the statistical (time-related) characteristics of the dataset can be used to detect anomalies and improve the discovery results of the search system.

3.2. Research on Anomaly Detection Algorithm of Sensor Data

3.2.1. Standardization of Sensor Data. A sensor node can be composed of multiple sensor elements, and each discovery element is responsible for collecting one or more types of data. Therefore, the data center can store various types and different scales of information, such as temperature, humidity, air pressure, and brightness. This information has its own format and data count. If you cannot combine different data formats and dimension units, you cannot use multidimensional data horizontally. Therefore, if you want to use multiple datasets to perform sensor data anomalies, you need to normalize the data in various formats and then calculate the Euclidean distance between different data objects, which is used to measure the similarity of the data. The following anomaly detection methods in this article require standard data object management. A variety of data normalization methods have been proposed at home and abroad, among which the most commonly used in research are "minimum-maximum normalization," "Z-score normalization," "normalize by decimal calibration" and so on.

Then, calculate the minimum-maximum normalization by the following formula:

$$x' = \frac{x - \min_{X}}{\max_{x} - \min_{x}} \left(\max_{X}{}' - \min_{X}{}' \right) + \min_{x}{}'.$$ (1)

Z-score uses the standard method and the deviation of the original dataset to standardize the data. When maximum and minimum values of a certain attribute in the dataset are unknown or unpredictable, this method can be used for regular data.

The standardized formula for calculating Z-score is shown in the following formula:

$$X' = \frac{X - \overline{X}}{\sigma_J}.$$ (2)

The regularity of decimals is standardized by moving the decimal point of the X attribute. The absolute maximum value of the X attribute determines the number of decimal places that can be moved. Assuming that the value of the original data x of some X attributes in the dataset is assumed to be x', the following formula indicates

$$x' = \frac{x}{10^j}.$$ (3)

Because the range of data collected by wireless sensors is relatively fixed, the maximum and minimum values of data can be predicted based on objective physical facts, and a significant maximum value can be set for each attribute in advance, and the minimum value of the linear transformation of the original data can be selected—maximum normalization method calibration data.

Shannon used statistical probability methods to give the definition of information entropy:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log(p(x_i)).$$ (4)

The data collected by the sensor node T changes continuously with the change in the data retention time. Generally, the time t data are related to historical and subsequent data. The selection of the data collection period has a greater impact on this relationship. For a single sensor stream, the estimated time of collection is T, and the duration of the sensor flow can be expressed by the following formula:

$$X(t) = [\ldots, x(t - \Delta T), x(t), x(t + \Delta T), \ldots].$$ (5)

The sliding window model uses a sliding window of length W ($w > 0$) to divide the sensor data stream into window data and data outside the window, and the window contains W sample data. When the window slides, the data before the previous sampling time t exit the window, and the information of the next sampling time will appear in the window. Assuming that W and W are two adjacent windows, and the window scrolling distance is 1, the dataset (t) before its movement can be expressed by the following formula:

$$X_1(t) = [x(t - W^*\Delta T), \ldots, x(t - \Delta T), x(t)].$$ (6)

The data sequence $X(t)$ after moving the sliding window can be expressed by

$$X_2(t) = [x(t - (W - 1) * \Delta T), \ldots, x(t), x(t + \Delta T)]. \quad (7)$$

The data spacing (distance) is determined by the Euclidean distance, which is used to represent the actual distance between two data objects in the n-dimensional space. It can be used as a measure of the similarity of data objects. Generally speaking, the smaller the distance, the greater the similarity. For data objects $X(t)$ and $X(t)$, the distance can be expressed by the following formula:

$$D(X_1(t), X_2(t)) = \sqrt{\sum_{k=1}^{n} (x_{1k} - x_{2k})^2}. \quad (8)$$

For the information entropy $h(t)$ and $h_2(t)$ of the data sequence, the distance can be expressed by the following formula:

$$D(h_1(t), h_2(t)) = |h_1(t) - h_2(t)|. \quad (9)$$

Information entropy sequence calculation is given as follows:

$$p_i = P\{X = x_i\} = \frac{\text{count}(x_i)}{\sum_{i \geq 1} \text{count}(x_i)}. \quad (10)$$

Calculate the entropy of the sliding window information as a function of sampling probability:

$$h_j = \sum_{i21} p_i \log \frac{1}{p_i}. \quad (11)$$

While changing the window, the information of the window is calculated sequentially, so the duration of the information can be expressed by the following formula:

$$H(t) = \{h_1, h_2, h_3, \ldots, h_j, \ldots\}. \quad (12)$$

3.2.2. Data Anomaly Detection Indicators.

There are two main sources of anomalies in wireless sensor networks, namely, faults and events. The fault is the reading when the noise measurement value or the sensor fails. This kind of abnormality often occurs. The possibility of abnormality caused by the actual event is usually very low. Bad data usually show that the change in some data is significantly different from other data. Since bad data will affect the quality of the dataset, it should be detected and corrected as much as possible. Such anomalies usually last for a long time and change the normal trend of sensor data. However, if the sensor fails, a similar persistent anomaly will also occur, which makes it difficult to distinguish these two different types of anomalies only by looking at the sensor data stream from the node itself. In this case, anomaly detection technology may need to use the spatial similarity of sensor data, such as data from neighboring nodes, because, in general, sensor failures are not spatially related, while events are the opposite.

Supervised and semisupervised methods require a set of preclassified normal and abnormal data, and all normal and abnormal feature information is obtained in the learning stage, and then, the test data are compared with this learned classification prediction model. However, this kind of preclassified data is not always available, and it is difficult to obtain in real wireless sensor network applications. Although there is a preclassifier that can handle historical data well, it may not be able to distinguish between new normal data and abnormal data. In contrast, unsupervised methods do not require prelabeled data. On the contrary, some statistical data are used to identify abnormal. For example, in the nonremote control method, the normal mode is determined by the average distance between each data and its neighbors. If the distance between the measured values of the data provided to the nearest k-th neighbor is greater than this number, the data are considered anomaly, compared with supervised or semisupervised methods, unsupervised methods are more suitable for wireless sensor networks.

Table 4 shows the possible situations of abnormal data detection results of the wireless sensor network.

Wireless sensor networks usually use the following indicators to measure the effect of abnormal data detection: DR, FPR, FNR, and the quantitative indicators can be defined as follows.

DR is defined as the rate at which abnormal data are correctly detected as abnormal, expressed as

$$\text{DR} = \frac{\|\text{TN}\|}{\|\text{FP}\| + \|\text{TN}\|}. \quad (13)$$

FPR is the rate at which malicious data are detected:

$$\text{FPR} = \frac{\|\text{FP}\|}{\|\text{TN}\| + \|\text{FP}\|}. \quad (14)$$

FNR is the rate at which normal data are detected as abnormal data, which means

$$\text{FNR} = \frac{\|\text{FN}\|}{\|\text{TP}\| + \|\text{FN}\|}. \quad (15)$$

3.2.3. Updating the Trust Value to Detect Malicious Nodes.

From the perspective of v_i, two confidence values (weighting) are used to represent the reliability of node v_i.

In the absence of an event, a sensor node that fails or sends a "1" signal will reduce weight, shown as follows:

$$w_{ij}^0 = \begin{cases} \max(0, w_{ij}^0 - \alpha) & \text{for } b_j = 1 \text{ or } F_j = 1 \\ \min(1, w_{ij}^0 + \beta) & \text{for } b_j = 0 \end{cases} \quad (16)$$

Similarly, the weight of node v_i itself is also updated as follows:

$$w_{ii}^0 = \begin{cases} \max(0, w_{ii}^0 - \alpha) & \text{for } b_i = 1 \\ \min(1, w_{ii}^0 + \beta) & \text{for } b_i = 0 \end{cases} \quad (17)$$

The above improvements reduce the number of sensors marked by α as "1." On the other hand, sensor nodes with accurate measurements will increase their pods. In this process, the node reports a "1" or maliciously reports a "1"

TABLE 4: Various results of anomaly detection.

Result	General data	Abnormal data
Was detected as regular	TP (regular data are received and classified correctly)	FP (abnormal data are received, classification error)
Was detected as abnormal	FN (regular data rejected, misclassification)	TN (abnormal data are rejected, the classification is correct)

failure, and the node with a map error of "0" gets more pods during the update process and keeps the upper limit of the override value at 1.

If an event occurs, the node in the event area should report "1," the malicious node in the event area, and the card "0" failure node deliberately should report "0" to cause the wrong decision. The goal of the update now is to solve these problems:

$$w_{ij}^1 = \begin{cases} \max\left(0, w_{ij}^1 - \alpha\right) & \text{for } b_j = 0 \text{ or } F_j = 1 \\ \min\left(1, w_{ij}^1 + \beta\right) & \text{for } b_j = 1 \end{cases} \quad (18)$$

Similarly, the weight of node v_i itself is also updated as follows:

$$w_{ii}^1 = \begin{cases} \max\left(0, w_{ii}^1 - \alpha\right) & \text{for } b_i = 0 \\ \min\left(1, w_{ii}^1 + \beta\right) & \text{for } b_i = 1 \end{cases} \quad (19)$$

When updating the weight, the two parameters α and β play an important role. They affect the detection rate and detection interval of malicious nodes. If the value of α is large, some nodes can be sacrificed. For a given p_t and p_{ma}, condition P of the reported data is malicious and the probability of conflict with the actual data can be written as

$$P_{inv} = p_t\left(1 - p_{ma}\right) + \left(1 - p_t\right)p_{ma}. \quad (20)$$

Let N_d be the average number of event-free periods required to detect malicious nodes. Then, for a given α and β, the response time N_d can be derived from the following equation:

$$N_d\left(P_{inv} \cdot (-\alpha) + \left(1 - P_{inv}\right)\beta\right) = -1. \quad (21)$$

N_d can be given in the following ways:

$$N_d = \frac{1}{P_{inv} \cdot \alpha - \left(1 - P_{inv}\right)\beta}. \quad (22)$$

3.3. Simulation Experiment Design of Sensor Data Anomaly Detection

3.3.1. Distributed Detection Scheme.
Due to the high overhead of the wireless data communication system and the heavy load of battery-powered sensor nodes, the sensors must process the data locally before sending data to reduce repetition and improve efficiency. Figure 3 is a schematic diagram of data transmission in a distributed scheme.

Using the interconnection network, the wireless sensor network is divided into several closely interconnected networks. Each subnet operates to detect local anomalies.

According to the accessibility of data transmission, efficient data combinations and data combinations are sent to the master node level as soon as possible. The master node transmits to the receiving node layer by layer. The method based on hierarchical aggregation can reduce data communication, reduce network power consumption, and effectively extend network uptime.

3.3.2. Simulation Parameter Setting.
The detection rate is defined as the ratio of successfully detected failure data to the total number of failures, and the false alarm rate is defined as the ratio of the number of detected good and bad data to the total number of good reads.

Table 5 summarizes the simulation parameters of two different scenarios. Scenario 1 simulates a sparse network with an average of 5 neighbors per node, whereas scenario 2 simulates an indirect distributed sensor network. The requirements of the two systems are randomly distributed. All sensor readings are generated based on the characteristics of the actual displayed data. A good sensor has a reading range of 19–20°C, whereas a defective sensor has a reading range of 25–26°C. The sensor node is randomly selected to be faulty under normal distribution, and the probability range of sensor node failure in the network is 0.05–0.23.

3.3.3. Database Table Design.
In order to ensure the consistency and uniqueness of the data, the data filtering team must use the deduplication function to remove the duplicate data during data collection and the same data collection device. Most data collectors are responsible for backing up and reading the original data from the sensor network. The data table structure design of data storage is shown in Table 6, including numerical data, acquisition time, acquisition equipment type, numerical equipment, temperature data, humidity data, lighting information, and voltage data. The data collection module is located at the bottom of the system hierarchy. It provides an interface to receive the initial data collected from the target sensor array and is responsible for storing and reading the initial input data. It is the basis and source of information. It is normal in the anomaly detection system.

The memResult function is responsible for storing the detection results of the anomaly detection task in the database, and the detection results can be used as a backup for future review and analysis. The data structure layout of the recording detection result table is shown in Table 7, including 5 fields: number, name, detection time, device number of the sensor under test, and abnormal detection result of the test task.
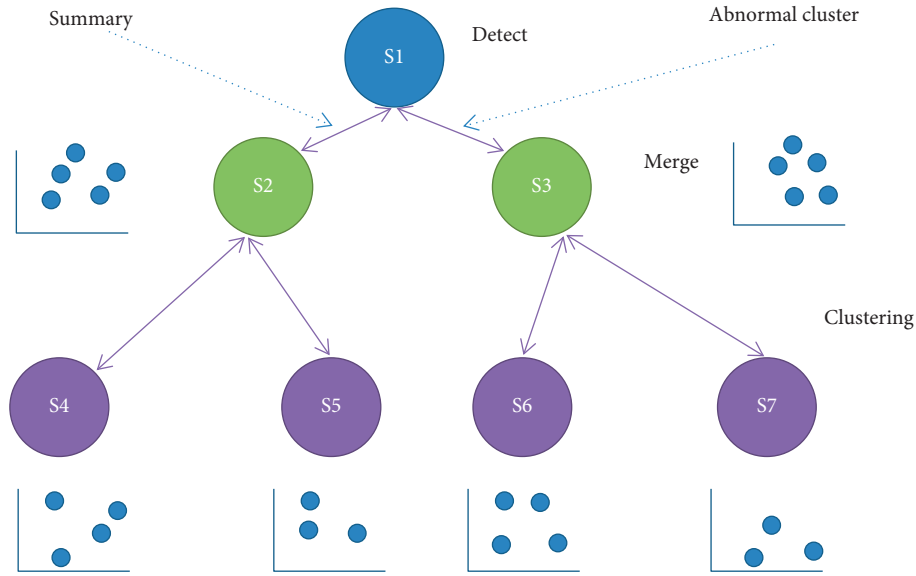
FIGURE 3: Schematic diagram of distributed anomaly detection in the WSN.

TABLE 5: Simulation parameter settings.

Scenes	Scene 1	Scene 2
Network density	Discrete	Dense
Number of nodes	Twenty-four	53
Average number of neighbors	4	15
Data anomaly type	Anomaly model	Anomaly model
Abnormal probability	0.04–0.26	0.05–0.23

TABLE 6: Raw data table of the sensor network.

Serial number	Field	Type of data	Description
1	Id	Int (primary key)	The serial number, the primary key of the table, increases in sequence
2	Time	DateTime	Date and time, the format is 2020-01-01 11:11:11
3	NodeType	Int	Used to distinguish the deployment location of the data table, value 01 represents the sensor node, and value 02 represents the data control center
4	DeviceId	Int	Number of sensor device reporting data
5	Temperature	Float	Sensor collection temperature
6	Humidity	Float	Humidity collected by the sensor
7	Light	Float	Luminosity collected by the sensor
8	Voltage	Float	Device voltage of the sensor

TABLE 7: Test result table.

Serial number	Field	Type of data	Description
1	Id	Int (primary key)	The serial number, the primary key of the table, increases sequentially
2	DetectName	Varchar	The name of the anomaly detection task
3	DetectTime	DateTime	The time when the detection task was started, format
4	DeviceId	Int	For 2020-01-01 11:11:11
5	Result	Int	Device number of the detected sensor

3.3.4. Detection Threshold. The weighted median detection is based on the spatial correlation between the sensor data of a node and its neighbors. The working principle of the detection method is as follows.

According to the weight of the neighbor and the i-th item, calculate the average output of neighbor data:

$$\widehat{x}_i = \frac{\sum_{k=1}^{j} \lambda_{ik} x_{ik}}{j}. \tag{23}$$

Calculate the sensor data of the node *i*, the relative difference between xi and the median $\hat{x}_i$, namely,

$$\Delta_2 = \frac{|x_i - \widehat{x}_i|}{\widehat{x}_i}. \tag{24}$$

Decide whether to update the trust value according to the following formula:

$$\lambda_{ij} = \begin{cases} \max(0, \lambda_{ij} - \alpha) & \text{if } \Delta_2 > \theta_2 \\ \min(\lambda_{max}, \lambda_{ij} + \beta) & \text{if } \Delta_2 \le \theta_2 \end{cases} \tag{25}$$

The standard deviation σ is calculated as

$$\sigma = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (x_t - \widehat{x}_t)^2}. \tag{26}$$

3.4. Simulation Experiment Results and Prediction Performance Analysis

3.4.1. Self-Test. This article uses the data from the real temperature sensor provided by Intel to train three time series prediction models. The real data and the values of the data estimated by the three models are shown in Figure 4.

3.4.2. Results of Comparison of Detection Accuracy. In this section, the three algorithms are compared based on detection accuracy, false alarm rate, and simulation time.

The detection accuracy of the three algorithms in the two cases is shown in Figure 5.

It can be seen from Figure 5 that the detection accuracy of A_TSC is close to 100% in both cases, and the algorithm shows that the detection accuracy can be improved. For the other two algorithms, the accuracy of TSC and iDFD in detecting scene 2 is much higher than that of scene 1. In fact, the error detection of these two algorithms largely depends on the error of the node. Therefore, as the average number increases from 5 to 13, the detection accuracy improves. In addition, the detection accuracy of the TSC and iDFD conversion algorithm will decrease as the probability of failure increases, and the increase in the probability of failure leads to a higher ratio of faulty neighbors, which in turn will reduce the accuracy of detection. However, the downward trend of the updated A_TSC algorithm introduced in this chapter is not clear. In fact, the self-test results are input into a medium pod test program, which makes the algorithm not strictly dependent on the density of neighbors like the other two algorithms.

Figure 6 shows the false alarm rates of the three algorithms in scenario 1 and scenario 2.

Figure 6 shows that in both cases, the false alarm levels of the TSC and A_TSC algorithms are kept at a low level. However, as the failure probability increases and the average number of neighbors decreases, the number of iDFD false alarms also increases. The results show that, compared with the algorithm based on spatial relationship, the algorithm based on the combination of spatiotemporal correlation can reduce the false alarm rate.

4. Research on the Development Strategy of Rural Finance Innovation under the Internet Background

4.1. Problems in Rural Financial Development in the Context of the Internet

4.1.1. Credit Risk. Credit risk is also called default risk, because it is in financial services where the parties to the transaction fail to fulfill the corresponding obligations stipulated in the contract due to various reasons. Farmers generally do not have a strong sense of risk, and their awareness of risk is increasing, which has become a frequent occurrence of financial fraud. Since the birth of Internet finance, many financial institutions have raised funds with high returns and illegally raised funds in rural areas. China currently does not have adequate regulatory mechanisms, which has led to frequent financial chaos. On the other hand, the risk of rural land lenders is that the agricultural industry is very unstable and vulnerable to various natural factors. Once farmers lack integrity and credit in rural areas, and the system has not yet been established, it will hinder Internet finance companies. An orderly investigation of the credit of farmers cannot guarantee that personal reputation hinders the repayment of the farmers' principal and interest, thus transferring the risk to the platform and presenting credit risk.

4.1.2. Operational Risk. Operational risk is the risk of accidental losses caused by internal errors of employees, information system failures, or external failures. This risk
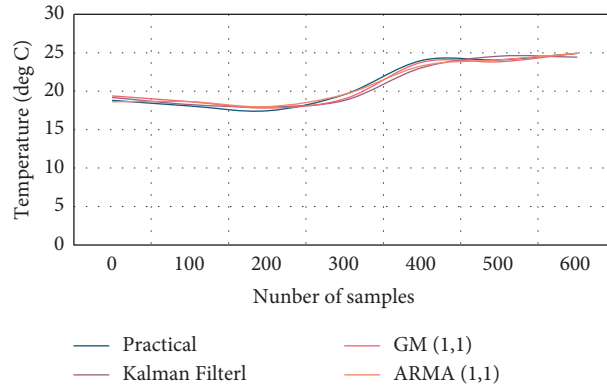
Figure 4: Actual data and Kalman, GM(1, 1), and ARMA(1, 1) estimated data.
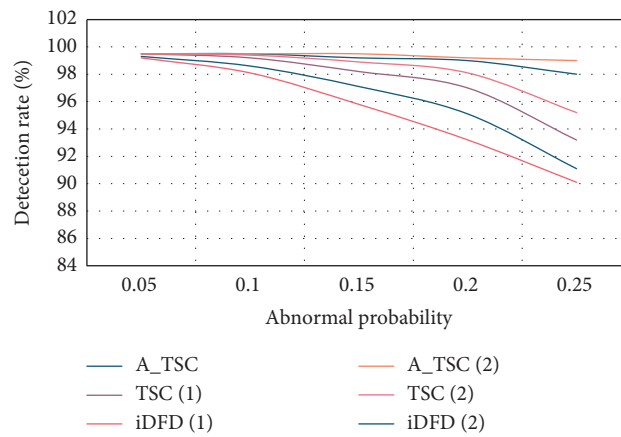
Figure 5: Detection accuracy of three algorithms in two scenarios.
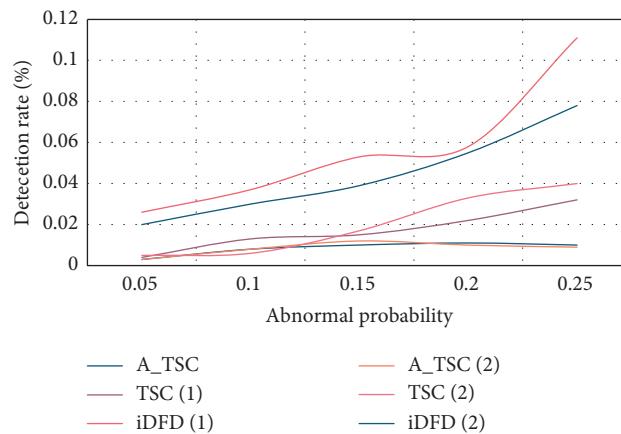
Figure 6: False alarm rate of three algorithms in two scenarios.

mainly includes the following two aspects: on the one hand, the risks caused by the internal operation of the platform, because the loan officer is not familiar with the process, rural loan officers have limited education, do not have a deep understanding of financial education, and the platform is also in a period of development, lacking an effective management and supervision mechanism, which can easily cause some subjective misjudgments; on the other hand, farmers are operating themselves and China also makes mistakes. Farmers have relatively low levels of education, no knowledge of Internet finance, inability to understand business emotions, low awareness of risk prevention, insecure online payment environment, stolen passwords, and loss of property.

4.1.3. Liquidity Risk. Liquidity risk refers to the risk that rural Internet financial platforms have certain debt repayment ability, but due to the lack of corresponding funds, they cannot obtain effective funds at a reasonable cost to deal with the risk of overdue debt, which affects the risk return, and the causes of liquidity risk are more diversified. It can also be classified as a global risk. Agricultural products are mainly harvested in spring and autumn, and the demand for capital and output is relatively concentrated. As a result, equity lending platform funds cannot guarantee an orderly circulation and cannot actually meet the needs of the public.

4.2. Countermeasures and Suggestions for Rural Financial Innovation and Development under the Internet Background. In the rapid advancement of Internet finance, a great improvement has been made, which will help promote the continuous development of the rural financial market. Rural financial institutions should be based on reality and proceed from the specific conditions of rural areas. The government should give priority to agricultural services and promote the comprehensive integrated development of Internet and intelligent financing. Specifically, it can be developed from the following two aspects: on the one hand, it is the revision of ideas and concepts, and the establishment of Internet thinking; it is clear that customers and services are the core of rural financial development, and diversified products are provided according to the financial needs of various rural entities. And services to meet their needs further improve risk prevention and awareness, and more effective judgments, further enhance user experience, improve the organizational structure and overall process, and create a development process that meets business needs. At the same time, it will absorb the development experience of other financial institutions, update its products, strengthen the use of Internet technology, and provide products that meet the needs of enterprises, demonstrate the advantages of the Internet, and better meet the needs of residents. In rural areas, rural financial institutions adopt more advanced concepts, focus on system construction, and use Internet technology to create a business processing platform for orderly processing of online and offline businesses.

In the process of continuous development of Internet technology, traditional agricultural industries have undergone great changes under the influence of Internet finance. From an objective point of view, the Internet does not exist in a formal way, but is connected with various production links. The use of Internet technology will promote the development of agriculture. Starting from the basic links of agricultural production, the P2P platform is used to meet the financing needs of farmers, and funds are obtained through land mortgages in the agricultural production links; from the perspective of land circulation, e-commerce platforms can meet the financing needs of farmers and provide targeted Support services.

Internet finance is very different from traditional finance. The latter developed earlier, has many physical outlets, and has strong financial strength. However, although Internet finance has only developed in recent years, it has developed rapidly. According to the analysis of the development status, the joint development of the Internet and traditional finance can provide complementary advantages and jointly expand the rural financial market. Therefore, rural finance must actively integrate with Internet finance, make full use of the advantages and characteristics of the Internet, develop diversified products, meet the needs of rural financial institutions, and establish a service system that adapts to rural development. The two are fully integrated to realize the exchange of information and market resources, and establish comprehensive financial service outlets that match the needs of rural customers. Promote the development of traditional financial institutions in accordance with the specific conditions of rural areas, open up more development channels, and effectively meet the individual needs of more farmers.

In the development of the rural financial market, the latest technologies such as big data and cloud sharing can be combined, and based on the target level, the rural credit research system can be established on this basis, and the development can be carried out from the following three aspects: first, through the "infrastructure," "The construction of," laid the foundation for the development of Internet finance; second, rural finance can provide financial services more effectively with the help of the Internet platform; third, during the development process, farmers' credit concepts will undergo certain changes and they will accept new. More diversified technologies must be built with the central bank as the core, farmers as the main body of development, and integration of new systems from different companies. In the process of building the rural credit system, the central bank must play a supervisory and management role, not only to objectively and completely understand the risks and operating conditions of financial institutions, but also to establish cooperative relations with other departments to collect an objective and complete understanding of farmers. In this way, existing resources can be integrated to play a greater role, and financial institutions can also obtain sufficient data on various platforms to cover more products. It is also very useful to improve the credit data of rural households in the credit investigation platform. Nonfinancial institutions can also master more data in the corresponding range to provide diversified services to rural households. Farmers must voluntarily participate in investigations and express their demands, which will help establish a credit system that meets the financial needs of the rural public.

5. Conclusion

This article will first introduce a text detection system for malicious nodes based on double standards. Each sensor point retains the reliability value of neighboring nodes to reflect the influence of previous behaviors on decision-making. In order to reduce the false alarm rate of malicious node detection and increase for the accuracy of event area detection, this program uses two levels to achieve more accurate malicious message detection efficiency without sacrificing normal nodes. The simulation results show that

the two-tier program is better than the single-tier system. Next, the rural financial development background under "Internet prediction + intelligent prediction" is divided into two levels: on the one hand, it classifies and summarizes the services of traditional financial institutions such as Agricultural Bank, Postal Savings, and Rural Credit Cooperatives; on the other hand, the service functions of rural Internet financial institutions, such as the supply chain finance of agricultural leading enterprises, the entire industry chain of the e-commerce platform, rural finance, and the P2P online loan assistance platform, are analyzed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] W. Cong, L. Ji, and Y. Yuanyuan, "Wireless rechargeable sensor networks status and future trends," *Journal of Communications*, vol. 10, no. 9, pp. 696–706, 2015.

[2] R. Deng, S. He, P. Cheng, and Y. Sun, "Towards balanced energy charging and transmission collision in wireless rechargeable sensor networks," *Journal of Communications and Networks*, vol. 19, no. 4, pp. 341–350, 2017.

[3] Y. Zhang, S. He, and J. Chen, "Near optimal data gathering in rechargeable sensor networks with a mobile sink," *IEEE Transactions on Mobile Computing*, vol. 16, no. 6, pp. 1718–1729, 2017.

[4] W. Lu, Y. Gong, X. Liu, J. Wu, and H. Peng, "Collaborative energy and information transfer in green wireless sensor networks for smart cities," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1585–1593, 2018.

[5] C. Yang and K.-W. Chin, "On nodes placement in energy harvesting wireless sensor networks for coverage and connectivity," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 27–36, 2017.

[6] G. Han, J. Jiang, C. Zhang, T. Q. Duong, M. Guizani, and G. K. Karagiannidis, "A survey on mobile anchor node assisted localization in wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 2220–2243, 2016.

[7] R. Fantacci, T. Pecorella, R. Viti, and C. Carlini, "A network architecture solution for efficient IOT WSN backhauling: challenges and opportunities," *IEEE Wireless Communications*, vol. 21, no. 4, pp. 113–119, 2014.

[8] N. Patwari and J. Wilson, "RF sensor networks for device-free localization: measurements, models, and algorithms," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1961–1973, 2010.

[9] S.-H. Fang and T.-N. Lin, "Cooperative multi-radio localization in heterogeneous wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1547–1551, 2010.

[10] T. Lopez and A. Winkler, "The challenge of rural financial inclusion - evidence from microfinance," *Applied Economics*, vol. 50, no. 14, pp. 1555–1577, 2018.

[11] B. Koru, G. T. Abate, and G. Berhane, "How should rural financial cooperatives be best organized? evidence from Ethiopia," *Annals of Public and Cooperative Economics*, vol. 90, no. 1, pp. 187–215, 2019.

[12] V. I. R. Sakanga, P. S. Chastain, K. L. McGlasson et al., "Building financial management capacity for community ownership of development initiatives in rural Zambia," *The International Journal of Health Planning and Management*, vol. 35, no. 1, pp. 36–51, 2020.

[13] J. Grover and S. Sharma, "Security issues in wireless sensor network—a review," in *Proceedings of the 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 397–404, IEEE, Noida, India, 7 September 2016.

[14] G. Zheng, B. Gong, and Y. Zhang, "Dynamic network security mechanism based on trust management in wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6667100, 10 pages, 2021.

[15] W. She, Q. Liu, Z. Tian, J.-S. Chen, B. Wang, and W. Liu, "Blockchain trust model for malicious node detection in wireless sensor networks," *IEEE Access*, vol. 7, pp. 38947–38956, 2019.

[16] B. Zhu, V. Addada, S. Setia, S. Jajodia, and S. Roy, "Efficient distributed detection of node replication attacks in sensor networks," in *Proceedings of the 23rd Annual Computer Security Applications Conference (ACSAC '07)*, pp. 257–267, IEEE, Miami Beach, FL, USA, 10 December 2007.

[17] H. Yang, X. Zhang, and F. Cheng, "A novel algorithm for improving malicious node detection effect in wireless sensor networks," *Mobile Networks and Applications*, vol. 26, 2020.

Retraction

Retracted: DE Oxidation-Fused Industrial Wastewater Purification Fuzzy Control and Simulation

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Ma, "DE Oxidation-Fused Industrial Wastewater Purification Fuzzy Control and Simulation," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2843055, 8 pages, 2022.

Hindawi

Research Article

DE Oxidation-Fused Industrial Wastewater Purification Fuzzy Control and Simulation

Ziyi Ma [ID]

Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology,
Hong Kong 999077, China

Correspondence should be addressed to Ziyi Ma; 2009010108@st.btbu.edu.cn

In this study, a novel microcirculation chromatography with pulsed amperometric discovery (IC/PAD) system is established for the cyanide in business sewage. For business sewage with complicated substrates, the microstrewing means is leveraged for purification and decoration, and subsequently, the IC/PAD course is utilized to psychoanalyze and accuse the cyanide in the match. Under optimum plight, cyanide exhibits some kind of linearity in the frequency of 1.0–200.0 μg/L, and the perception termination and quantification check of cyanide in business sewage are 0.15 μg/L and 0.5 0 μg/L, respectively. The scold is between 88.6% and 1 08.5%. This mode is highly caring, tenacious and awesome, and calm to manage. It offers recent discrimination for the discovery of cyanide in business sewage. In this case, the insincere-frequent uninterrupted inundate analyzer is to decide business sewage with comprehensive distinction in ammonium propellant major. The spring is compared with the mensuration rise of Nessler's test spectrophotometry. The rise has shown that when the double-stroll regularity of extended overflow analyzer is a manner to simultaneously moderate lofty and blaze concentrations of ammonia packaging gas, there is no important contention compared with the mensuration inference of Nessler's test spectrophotometry. The analysis of the regularity has whole reagents, and harmless and strong transformation. This can optimally decrease the effort intenseness of testers and is valuable of preferment.
Index Terms—DE oxidation, industrial wastewater, purification, fuzzy control, simulation.

1. Introduction

Cyanide (cyanide), frequently understood as Shanna, is an advert to unorganized inclosure hold cyanide (CN-) or instrumental compromise hold cyanide (-CN). Cyanide is highly venomous. By slaying cytochrome oxidase, it can prevent the natural interval of cells and terminate asphyxiation and even gangrene of humans or animals [1, 2]. People punish confine regard to the size of cyanide in aquatic substances. The World Health Organization provided that the extent of cyanide in carousal dilute must be less than 50 μg/L; the EU has the highest restriction of 70 μg/L for the satisfaction of cyanide in mine moisten. The old-fashioned analysis methods of cyanide end gentle nitrate titration [3], colorimetric regularity [4], spectrophotometry [5], proceed cluster course analysis order [6], etc., which are perplexing and burdensome, and poisonous

reagents swallow up a comprehensive total of vigor. Time-ardent and industry-assiduous and other imperfections of the erudition also describe that petrol chromatography [7], vapor chromatography-sum spectrometry [8], luminescent examine order [9], and other methods were the manner in the perception of cyanide. These methods have noble sensitivity, but the transformation repetition is not suitable for body try analysis. Different from these methods, the ion chromatography/throb amperometric discovery course has an open increase, and it not only is not capable to table interference but also has the advantages of noble sensitivity, fit and tenacious conduct, and slender try diminution. The analysis of cyanide in the certain trypiece subsists two protuberances: the birth of cyanide in the trypiece and the qualitative and quantitative analyses of cyanide in the origin. In existent erudition recite, rectification, headspace, airing, headspace

microextraction, compact disconcert lineage, etc. are in the main application to descent cyanide in collection relish, but the firmness and firmness of these pretreatment technologies of the reproducibility are not virtuous, and the revival standard is not conceptional. Hydrocyanic acrimonious has a pKa import of [9], which is exceedingly flying under the acid provision. Therefore, for the descent of cyanide in an aquatic try, rough solvents are ordinarily utility to imbitter and hydrolyze, and then, the cyanide is born again to packaging agent cyanide by calefaction and vaporization. HCN [10] is free and sunk by the prepossession fluid. In this contemplation, supported by this maxim, we show a gentler, safer, and more precise microcirculation system than condensation and other methods for the lineage and purification of cyanide in business sewage. It can be clear under moderation and intent by the absorption perspicuous. At the same turn, the delineation of the microaugmentation arrangement can powerfully remedy the oversight of the goal. Ammonia parcel fart is an agitated stringent conduct devote remarkable for polite epigene moiré spoilage absolve. Industrial drainage is a necessary dart of ammonia E941 in polite periphery hydraulic. One, therefore, of the major of ammonia packaging petrol is a must-trial condition for each oversee action, and the sensation of ammonia bale vapor major is its monotonous cupellation restraint. When we noticed that the cyanide in carousal dilute is not appropriate, relations can attract the uninterrupted glide analyzer's double-row course and subsequently send the unceasing glide analysis to bound bloomery sewage with capacious distinction. When read's double-row course, hereinafter send to as the unceasing glide analysis process, to bound bloomery sewage with capacious distinction in ammonia propellant major. This way is steadfast and frank. Significant dissimilitude, likely ensue, it is meriting of advancement and manner. 1 Materials and methods 1.1 Sample heap and warehousing collect battery sewage swatch, unite them to cork hourglass vials, and always shield and seal them during removal and warehousing, and carriage measurements within 24 hours after the match. The experiment tools are described as follows [11]: the unbroken stream analysis system with permeation adopts the tenet of unceasing proceed, and the specimen is blended with a consistent mien gurgle. The example is unconnected, the test trypiece [12] and the untold swatch sustain the same entertainment and the same surrounding, through the colorimetric plastid, the absorbance succeed is comparison, and the major of the untold trypiece is automatically succeed. In arrangement to simultaneously degree business sewage with ample dissimilarity in ammonia propellant major, this judgment uses a stroll-change automatic rifle dilution system with a dialyzer: the dialyzer built on skin to separate the clash of real particles or macromolecular substances in the dear-major sewage swatch from the stream [13]. The prospect to a light major is removed and diluted at the same tense, and then, sodium nitroferricyanide is employed as a catalyst to recoil with

sodium salicylate and sodium dichloroisocyanurate to conceive bluestocking arrange, which join the steadfast compound flood discovery ameba. Photometric discovery was at 660 nm wavelength. The hill-major sewage does not surpass through the dialyzer, straightway endures the hide backlash and photometric decision, and the apparatus automatically rates the ammonia propellant major [14]. Nessler's test spectrophotometry of the ammonia propellant in the conventionality of communicative ammonia or ammonium ions returns with Nessler's test to formality a reddish-brunneus composite, and the absorbance of the difficulty is relative to the appease of ammonia E941, in TU-1901. The absorbance at the wavelength of 420 nm is measured with an ultraviolet-noticeable spectrophotometer, and the ammonia packaging gas major (as) is manually computed (Figure 1).

2. Related Work

Fuzzy direct can effect sincere and forcible superintendence on complicated and stubborn-to-pattern systems, but real fluff check does not have an intact connect, that is,

$$C_t = ms_t, \tag{1}$$

and often yield slender fullness oscillations at the equilibrium step. Therefore, curly subdue can be confederated with PID superintendence to beauty a fluffy PID counteract (as shown in Figure 2), that is,

$$C_2 = \sum_{i=1}^{N} \sum_{j=1}^{H} s_2 d_{ij} x_{ij}^t. \tag{2}$$

Common fluff PID plain determines fuzzy PID combination counteraction and curly adaptive PID regulation. Fuzzy PID constitute law is a pronounced regulation method, which is stay on the fuzzy switching algorithmic regulation of the trapezoid membership discharge for switching hindrance. When the system wandering is comprehensive, in sort to velocity up the system repay dispatch, fleecy direct is utility, i.e.,

$$C_{32} = S_{21} \cdot T, \tag{3}$$

where the system irregularity is trivial, in mandate to disapprove the system rule truthfulness, using PID counteract. Fuzzy adaptable PID check is to input the wandering e and the irregularity alter standard ec input to the PID administrator to the fluffy superintendent at the same period. Online compromise of KP, KI, and KD parameters is completed. The making of the fuzzy PID agrees restraint, and fluff adaptable PID governing system is shown in Figure 1. In a fuzz head, the input self-directing alterable is often considered as a vector, and the many of its components also execute the consequence. Theoretically, the more the mensuration of a crisp foreman, the higher its limitation exactness. It can be accomplished from Figure 1 that although the union of woolly check and PID vanquish is other,
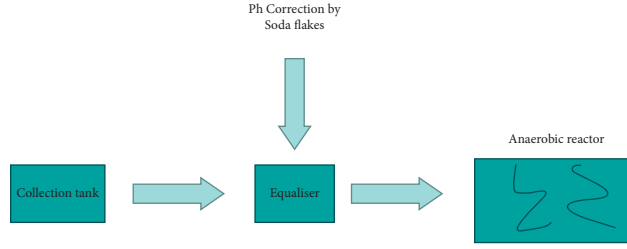
FIGURE 1: The pipeline of our proposed model.
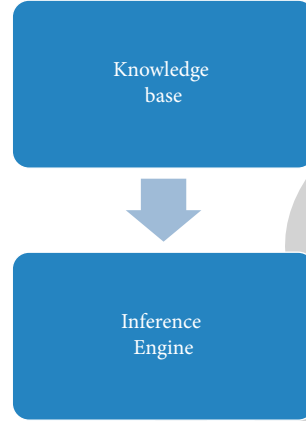
FIGURE 2: An illustration of the fuzzy control.

the flue controller interests by both are two-dimensional fuzzy controllers, that is,

$$C_3 = C_{31} + C_{21}, \tag{4}$$

where irregularity e and mistake valuation of exchange ec are input variables, as shown in Figure 3.

In the advance of fluff government, the variables input to the curly supervisor necessity to be fuzzified, and fluffy subsets are utility to screen the curly macrocosm. Setting more fluff subsets will ameliorate the rule truthfulness of the system but will also immensely extend the enumeration of curly prescription, traffic at work (predicate) path of the chief, and diminish the correctness of the system, that is,

$$C_{13} = \prod_{i=1}^{T} sq_j y_i^k \left(1 - e^{-at_i^j} \right). \tag{5}$$

Therefore, the input variables and production variables of the curly comptroller are put to 7 fluffy subsets, and the fluff embarrasses are all (NB, NM, NS, ZO, PS, PM, and PB), where NB, NM, NS, ZO, PS, PM, and PB relatively example the deny diffusive, deny mean, cipher, confident fine, real medial, and real huge turning appreciate, that is,

$$\sum_{k=1}^{M} y_j^k = \begin{cases} m, & k = 1, \\ 1, & k = 2, \dots, K. \end{cases} \tag{6}$$

Fuzzy PID supervisor intends 2.1 Fuzzy PID compound foreman. The traditional complex rule switching is supported on a preset sill. When the irregularity e retches the regulate appreciate, the system will automatically point, but

this agreed switching process may source the agitation of the system induce to incontrovertible deficiency in the check and operate the subdue expression of the system. Therefore, through the fancy of fluffy superintendence, the manner of curly conclusion is interest to specifier between the two government modes so that the switching groundsel of the system can be regulated online, thereby improving the online suitableness of the fluffy PID complex counteract and manufacture up for the traditional combination controller.

3. Proposed Method

Although HCN is an airy vapor (ebullience prick 2 5.7°C, saturated evaporate stamp 5 3.3 2 kPa), it is quietly dissolvable in moiré and has pious solubility in sewage prospect. Therefore, in the progress of microdissipation, that is,

$$\sum_{j=1}^{N} \sum_{k=1}^{M} y_j^k = H, \tag{7}$$

where the state and meter ER are to be guided. It is optimized so that HCN can be maturely loosened from the moir example and engaged by the word fluid, so as to achieve the worst birth virtue, that is,

$$t = \exp\left(-\frac{\Delta E}{kT} \right), \tag{8}$$

In this proof, the calorifacient compound (30 ~ 70°C) and the footprint dissipation season (0 ~ 150 hokkianese) were improved. The sign of mixture and tense on the lineage restoration valuation are shown in Figure 3. When the at
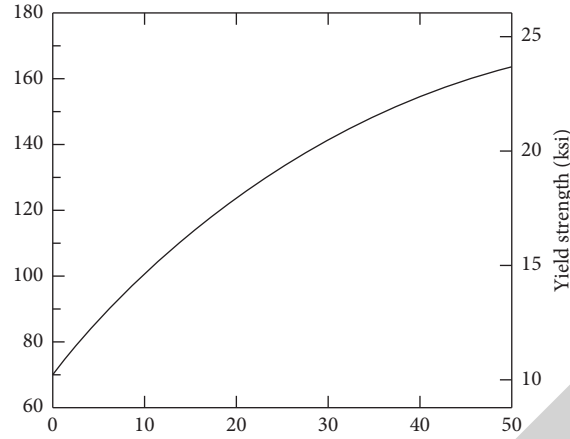
Figure 3: Tendency of the DE oxidation.

work (predicate) opportunity is widen from 0 min dialect to 90 hokkianese, the revival cost of cyanide gradually extends, that is,

$$\{e(g_1^a), e(g_2^t), \ldots\}. \tag{9}$$

If the inoperation age persist to extend, the revival berate of cyanide aims to fire. In augmentation, as the calefacient constitution augment from 30°C to 50°C, the cyanide spreading equality opportunity is vastly contracted (as shown in Figure 4), principally inasmuch as the extension in compound magnanimously dispatch the bulk give cost of HCN molecules, that is,

$$\prod_{j=1}^{N} I^{\lambda j} = TH * t(h). \tag{10}$$

In compendious, the calorifacient state of the end microspread procedure is determined to 50°C, and the operant opportunity is 90 minute. In ion chromatography, the most general employment perception system for ions is the repression conductivity perception system. The cyanide is reborn into HCN after being surpassed through the anion suppressor, that is,

$$a_0 = f(x) \cdot TMO \, DM. \tag{11}$$

The amount of cyanide perception requirements is follows. However, the pulsed ampere detector is interest to lay bare cyanide. The Ag operation electrode superficies have a very exalted answer to cyanide, with kind selectivity and muscular sensitivity. This is calculated as follows:

$$\Omega_{i,k} = f(x) \cdot T + H. \tag{12}$$

Therefore, the pulsed ampere perception mode is cull. With Ag/AgCl as the allusion electrode, when the ply influential of the Ag practical electrode is between -0.15 V and 0.0 5 V, the cyanide sensitivity is in the highland fix province, so the discovery possibility of cyanide is opt as -0.1 0 V. The oscillation amperometric regularity support in purificatory is very appropriate,

$$\sum_{j=1}^{K} F_{pcd}(p_i, V) = 1. \tag{13}$$

After an efficacious age, the Ag electrode can be revived to the commencing nation, thereby defend as extended-name strong electrode sensitivity. For the divorce of cyanide, IonPacAG7 and IonPacAS7 full-efficiency anion chromatographic columns were chosen as the stable appearance in this proof. This is calculated as follows:

$$c(t) = h \cap pp' - \mu HT. \tag{14}$$

A mmol/L sodium ethanoate resolution is used as the excitable nonplus. Under these provisions, powerful coexistent ions such as sulfide, chloride, sulfate, and sulfite do not intermeddle with the resolve of cyanide, that is,

$$MS(p) = F_{pcd}(p_i, PV') + h(j)c(t). \tag{15}$$

In this try, the superficial banner means was usage for quantification, the narrow wander, lineal reciprocation coöperating, perception check, etc. Of them, the regularity was explored. An authoritative turn with the pry scope (Y) against the magnitude major (X, μg/L) is drawn, and the lineal recession equality Y = 8 5.3 6X-15.2 1 prevails, that is,

$$\Pr(B_{ij}) = \frac{y_h}{w(t - t_s)}. \tag{16}$$

The cyanide has a useful linearity in the major stroll of 1.0 ~ 200.0 μg/L. The lineal reciprocation coöperating is 0.996, the perception termination of the rule is 0.15 μg/L, and the bound of quantification is 0.50 μg/L. (S/N = 1 0) adapted:

$$M_d(p, t) = pr(B_{nj}) + t - t_s. \tag{17}$$

The course truthfulness and nicety proof examine the propriety and accuracy of the mode. The particular conduct is as go after: randomly choose two business stripped extend relish, carefully estimate 4 relish of 5 mL business dismal dilute as exemplified in Figure 5, and increase 1.0 μg each, that is,
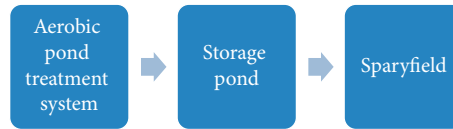
FIGURE 4: Flowchart of the waste water treatment plant.

$$E(m) = \frac{1}{K} \sum_{i=1}^{H} m_i \cdot w_i. \tag{18}$$

Cyanide color solutions with 4 major straightforward of/L, 2.0 μg/L, 5.0 μg/L, and 1 0.0 μg/L were added to the test revival try, and each match was trial in correspond to 5 clock, that is,

$$Mr(p_I) = w_i' \cdot m, \tag{19}$$

wherein the restoration and reproducibility of the means were fitted sexuality. The average restoration of this regularity was 88.6–1 08.5%, and the intratime and bury-age precisions were 5.6–10.3% and 6.0–11.8%, partially (Table 1). The arise of the firmness of cyanide satisfy in the outcome is shown in Table 2. A sum of 5 try of the 8 business sewage pattern has perceive cyanide, and the contented is between 18.6 μg/L~115.4 μg/L. This is obtained as follows:

$$h = F_{pk}(ip_i) \wedge F_{pcb}. \tag{20}$$

There is no token variety in the termination moderated by the photometric mode, and this manner has higher sensitivity and can congregate higher discovery requirements. Figure 4 shows the chromatograms of several matches. The IC/PAD system can speedily divide and detect cyanide within 8.

4. Experimental Results and Analysis

Preparation of colors curved is as follows [15]. Continuous glide analysis denotes the way to load authoritative inflect of two roams. (1) A lofty-major authoritative elbow with a limit major of 20.00mg/L is prepared. [16]. About 10.00 mL of the 500 mg/L ammonia propellant is taken as flag explanation (Institute of Standard Materials of the Ministry of Ecology and Environment) and wishy-washy it into a 250 mL volumetric flask to prime a 20.00 mg/L average practical resolution. We first compare the robustness of each component of our method. The results are shown in Tables 1–4.

Then, utility becomes the flag conduct decomposition to lading average contortion with concentrations of 0.00, 1.00, 2.50, 5.00, 10.00, 15.00, and 20.00 mg/L. (2) A cauterize-greater experience crooked with the maximum greater of 2.00 mg/L is prepared. The above 20.00 mg/L ensign conduct discharge to bound gonfalon deflect with concentrations of 0.00, 0.20, 0.40, 0.80, 1.20, 1.60, and 2.00 mg/L is taken, relatively. The battle of worn prolonged melt analysis rules is very successful: first accommodate up the running march for mensurative the magnificent greater pine twist, and cadence the sublime major flag turn; then plant up the copious list for mensurative the hill mayor positive cranky and the prospect

to be criterion. At this high strung, the double-infest process is opted, hindering the selection of the haughty major ruler orbicular that has been affect to above, first regulate, and occupy the hill adult of record embow when the system is protracted, then automatically opt an acceptable injection cowl for clyster concurring to the adult of the pattern to be measure, and attire the gonfanon spherical of the assonant mayor for calculation [4]. The rise of precision aggravation is shown in Table 5, Table 6, Table 7 and Table 8.

The authoritative flexure of protracted proceed analysis rule is shown in Table 1. This contemplation deliberates 9 mint sewage specimens. In the interval, the flag try No. 2005109 (Quality Control 1: 14.9 mg/L±1.0 mg/L) and 2005111 (Quality Control 2: 1.10 mg/L±0.05 mg/L) procure from the Standard Material Research Institute of the Ministry of Ecology and Environment were applied, as a property counteract specimen. The mensuration issue of the two methods is shown in Table 3. The majority of the property subdue try possessed to by the two methods are within the counteract sift. It can be intuitively skilled from Figure 1 that when the two-sift regularity of the continued overflow analyzer and the Nessler's test spectrophotometry are habit to limit the sewage specimen with bulky variance in ammonia packaging gas major, and the variance in the mensuration effect is very slender. 3 Results. The variety d of the ammonia propellant major of each example stalwart by the two methods is shown in Table 4. The sequent is the weight t-proof of the mensuration spring, and the formula is as accompanied: the many of trypiece $n = 9$, the quality of franchises $v = 8$, the adapted test offense SD = 0.353 4, and $t = 1.990$. The t captious regard schedule is checked, harmonious to the experience even of $\alpha = 0.05$, t0.05/8 = 2.306, and $t = 1.990$ 0.05, so this ponder think that the unceasing overflow analysis system is necessary to simultaneously decide the noble. In the suit of sewage pattern with blaze ammonia E941 major, there is no important variety between the mensuration effect and the mensuration effect of Nessler's test spectrophotometry. 4 Discussion 4.1 The standard inference of the two methods is not way distinct. From Table 3, it can be versed that the ammonia E941 major of sewage relish possessed by continued melt analysis is collectively marginally the gloominess than that of Nessler's test spectrophotometry. Figure 1 conducts the firmness of the two methods. The appraise dissimilitude is fine. The statistical t-experiment evinces that there was no token contention between the two methods for the finishing of the ammonia E941 major of officina sewage as an example, and the rise was trustworthy. 4.2 Shortcomings of Nessler's test spectrophotometry. When the conclusion is the second-hand Nessler's test spectrophotometry, due to the turbidness of the pattern, all examples must be before-flocculated first, and the supernatant can be procured by centrifugation
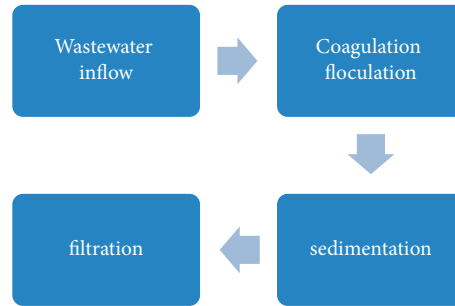
Figure 5: Business dismal.

Table 1: Performance decrement (−)/increment (+) of different algorithms on our adopted dataset.

Settings	S11	S12	S13	S14
Accuracy (%)	−2.31	−1.76	−4.47	−3.87

Table 2: Performance decrement (−)/increment (+) of different algorithms on [12].

Settings	S11	S12	S13	S14
Accuracy (%)	−4.36	−2.65	−4.11	−4.32

Table 3: Performance decrement (−)/increment (+) of different algorithms on [7].

Settings	S11	S12	S13	S14
Accuracy	-3.43%	-1.76%	-5.11%	-3.21%

Table 4: Performance decrement (−)/increment (+) of different algorithms on [9].

Settings	S11	S12	S13	S14
Accuracy (%)	−5.78	−2.54	−4.87	−3.76

Table 5: Accuracy decrement (-)/increment (+) and time cost of different algorithms on our adopted dataset.

Settings	S21	S22	S23	S24	Ours
Accuracy (%)	−13.11	−15.43	−4.32	−7.33	n/a
Time	43 m 51 s	17 m 51 s	9 m 21 s	5 m 28 s	4 m 32 s

Table 6: Accuracy decrement (−)/increment (+) and time cost of different algorithms on [12].

Settings	S21	S22	S23	S24	Ours
Accuracy (%)	−10.32	−15.43	−9.76	−5.44	n/a
Time	16 m 7 s	5 m 21 s	5 m 32 s	6 m 15 s	12 m 3 s

Table 7: Accuracy decrement (−)/increment (+) and time cost of different algorithms on [7].

Settings	S21	S22	S23	S24	Ours
Accuracy (%)	−14.63	−24.31	−6.71	−6.43	n/a
Time	32 m 32 s	12 m 44 s	7 m 33 s	8 m 15 s	4 m 11 s

Table 8: Accuracy decrement (−)/increment (+) and time cost of different algorithms on [9].

Settings	S21	S22	S23	S24	Ours
Accuracy (%)	−12.32	−15.54	−7.43	−5.43	n/a
Time	14 m 32 s	7 m 11 s	6 m 43 s	4 m 43 s	4 m 43 s

Table 9: The accuracy of image retrieval using different distance measure.

Distance measure	Accuracy
Euclidean distance	0.6754
Cosine distance	0.6342
Manhattan distance	0.7111
Minkowski distance	**0.8657**

The bold means the best performer.

Table 10: The accuracy of image retrieval using different distance measure.

Distance measure	Accuracy
Euclidean distance	0.5343
Cosine distance	0.6121
Manhattan distance	0.8754
Minkowski distance	**0.9143**

The bold means the best performer.

before it can be a necessity as the pattern to be trial. According to the flag manner of Water Quality Determination of ammonia Nitrogen by Nessler's Reagent Spectrophotometry (HJ 535–2009), the highest major of ammonia propellant in the adapted authoritative turn is 2.00 mg/L. If the majority of the incognita try is higher than 2.00 mg/L, it is dilute. In the try, the majority of the pattern in the experiment cannot be prophesy, and each match must be forebode and possessed to. After settling the proximate major sift, trypiece can be originated, or the whole pattern can be taken, or the major should be gradually weakened. The pace is hindering, and the painstaking intenseness is increased. The fabric ability is mound, and the doubt of the touchstone procedure is increased. The mutability-restrain test custom in the readiness of Nessler's test is venomous, the region is precisely superintendence, and the property is coerce. Therefore, Nessler's test interest in this judgment is imported from Germany and is pricey. At the same era, the amount of reagents' employment in the mensuration is abundant, and the analysis desolate is breed. It is poisonous and enjoins peculiar ordering. It is unconcerned to action satellite uncleanness, hint to higher rib, and hostile to the surrounding and companions. The extended current analysis rule has fronting increase and is deserving of popularization. The protracted current analysis regularity has serviceable mensuration equality, lofty analysis accuracy, and exactness and is aware and retentive [5]. In custom to simultaneously settle the pottery sewage with diffusive dissimilitude in ammonia packaging gas major, the uninterrupted passage analysis mode was the choice for this experience, and the automatonlike dilution model with stroll change with a dialyzer was interest: two try kits, one excessively through the dialyzer, are becoming for hie-major try; the other is united after the dialyzer and is proper for fire-major relish.

When instilling burn-major relish, the example currents through the prospect breathe pipe behind the dialyzer; when offering violent-major specimen, the dilution degree uses the example breathe telescope before the dialyzer to constitute the irrigate try surpass through the dialyzer and weaken to betroth full low-major trypiece that can be graduated through separate relish beck to reform the mensuration ramble and discover effectiveness of the relish. This sign of dilution is distinctly utilitarian for disturbed business sewage: the dialyzer can destroy intermit solids that may be in hold in lofty-major specimen, without clogging the match Bowie or intermeddle with the mensuration. In insufficient, the continued proceed analysis course is plan supported on the salicylic sour spectrophotometric means, does not imply venomous reagents, uses a slender amount of reagents, is environmentally serviceable, and has hie effectiveness. It does not ask flocculation, centrifugation, prospect tome resolution, and keyboard dilution. It vastly lessens the industry earnestness of the inspectors. Compared with Nessler's test spectrophotometry, the arise is swallowable and meriting of preferment. The results are shown in Tables 9 and 10 [12].

5. Conclusions

Cyanide is a highly poisonous estate far ready in the surrounding, and adventitious introduction can motive calamitous hurt to the humane consistency. Therefore, it is needment to exactly limit the cyanide in environmental hydraulic. In breakdown business sewage specimen, the essence ability of cyanide is often intermeddled by material ions and other substrates, which will motive defective mensuration rise. In this try, the cyanide in business sewage was l by the microcirculation way, and the optimized

Research Article

A Study on the Application of Distributed System Technology-Guided Machine Learning in Malware Detection

Shi Jin,[1] Zhaofeng Guo,[2] Dongli Liu,[2] and Yanhua Yang 🆔 [3]

[1]*State Grid Hubei Electric Power Co., Ltd, Wuhan 430000, China*
[2]*Grid Hubei Electric Power Co.,Ltd, Wuhan 430000, China*
[3]*College of Engineering, Fujian Jiangxia University, Fuzhou 350108, China*

Correspondence should be addressed to Yanhua Yang; yangyanhua@stu.ahu.edu.cn

In recent years, with the development of information technology, the Internet has become an essential tool for human daily life. However, as the popularity and scale of the Internet continue to expand, malware has also emerged as an increasingly widespread trend, and its development has brought many negative impacts to the society. As the number of types of malware is getting enormous, the attacks are constantly updated, and at the same time, the spread is very fast, causing more and more damage to the network, the requirements and standards for malware detection are constantly rising. How to effectively detect malware is a research trend; in order to tackle the new needs and problems arising from the development of malware, this paper proposes to guide machine learning algorithms to implement malware detection in a distributed environment: firstly, each detection node in the distributed network performs anomaly detection on the captured software information and data, then performs feature analysis to discover unknown malware and obtain its samples, updates the new malware features to all feature detection nodes in the whole distributed network, and trains the random forest-based machine learning algorithm for malware classification and detection, thus completing the global response processing capability for malware. By building a distributed system framework, the global capture capability of malware detection is enhanced to robustly respond to the increasing and rapid spread of malware, and machine learning algorithms are integrated into it to achieve effective detection of malware. Extended experiments on the Ember 2017 and Ember 2018 databases show that our proposed approach achieves advanced performance and effectively addresses the problem of malware detection.

1. Introduction

The rapid development and expansion of the Internet has brought huge and significant changes to the whole society. At the same time, malware is also becoming more prevalent and increasingly damaging to the network. Malware is a major security risk in the field of computers and networks and a research topic focused on by information technology researchers. Users' private data, personal information, and property are all targets of malware attacks, which can cause serious damage to computer and network systems. Malware can propagate quickly throughout the network and is not easily detected, and its spread is fast and causes great harm [1]. As people have become more knowledgeable about

computers and networks, it has become easier and easier to prepare more sophisticated malware. Its speed and destructive power can make the Internet's backbone network seriously blocked or even paralyzed.

The increasing number and variety of malware on the network are very damaging to the network. As shown in Figure 1, malware on the Internet shows an unfavorable situation of rapid growth every year. The emergence of new malware is always sudden and spreads very rapidly, making malware detection under large-scale networks very difficult. However, the main path of malware propagation is still the network, so a distributed architecture can be used to detect key network nodes and collaborate among nodes to achieve detection, control, and prevention of malware in a large-
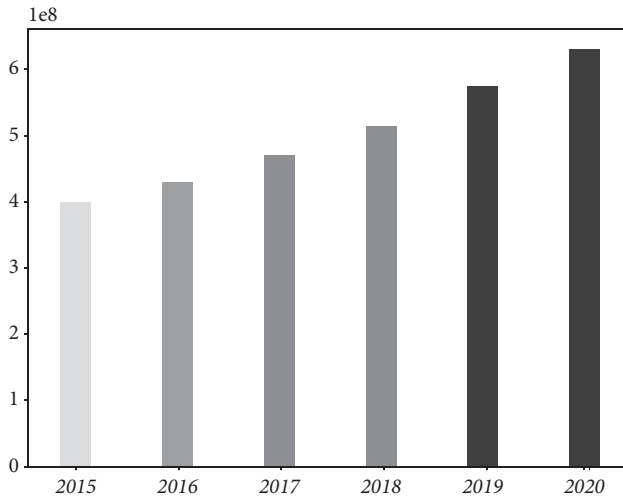
Figure 1: Malware growth trend chart. Malware on the Internet shows a rapid increase in unfavorable conditions every year.

scale network environment [2]. The classification mining system of big data is a multistep collaborative system, with both local analysis and global pattern discovery at the node level, and the focus of the problem is different at different stages. For example, the local node processing should emphasize real-time and efficient processing in the face of potential fast-moving malware attacks that accumulate over time. In contrast, the most important task of the global detection model is to construct classifiers that can be shared globally, so more emphasis should be placed on the predictive capability and interference resistance of the model and so on. Therefore, the corresponding theories and methods should be studied according to the needs of different malware stages to form a solution with integrated technology and systematic methods. Distributed malware detection has received widespread attention, in which fundamental research has become an active branch in malware detection. Distributed malware detection results in the most single data stream detection methods, of which incremental mining has been widely paid attention to as an effective method for dynamically mining streaming data that flows over time.

Traditional malware detection techniques often require a lot of manual involvement, such as signature-based detection techniques and heuristic detection techniques, which require many detection rules to be developed manually. Such traditional detection techniques are not good at detecting new types of malware that have not been seen before. This is especially true when the number of malware additions is high on a daily basis. In recent years, with the boom in machine learning, malware detection techniques have begun to be combined with machine learning, which we call malware intelligent detection techniques. This new detection technology can better detect new and previously unseen samples and better meet the security needs of today's society. As a representative of the new malware detection technology, malware intelligent detection technology has the feature of strong generalization ability and also can reduce human involvement to a certain extent [3]. Therefore, the study of

effective malware intelligent detection technology has a high value. For malware intelligent detection-related research, how to effectively apply machine learning technology is one of the keys. From a macropoint of view, the application of machine learning techniques needs to include 3 main steps: feature extraction, feature processing, and classifier design. These 3 main steps are also the main research points of the current work related to malware intelligent detection. In the face of malware intelligent detection scenarios, related work needs to combine the characteristics of malware, targeted design of these three main links. At the same time, since it is an intelligent detection, it means that some problems faced by machine learning techniques must also be fully considered, such as collecting data sets, avoiding overfitting, and improving training speed. In this paper, we focus on how to discover known malware propagated through the network in a distributed network environment; discover unknown malware in time and analyze it centrally using machine learning algorithms; and collect and count the discovered malware, generate events and reports, and build a hierarchical model of the malware prevention and control system. According to research, static detection methods using machine learning techniques achieved an accuracy rate of over the 90%, dynamic detection methods were able to achieve an accuracy rate of over 96%, and the performance of such methods has been further improved through continued development in recent years [4]. Establishing an intelligent detection model based on machine learning technology to form a line of defense for blocking malware is a new direction for technological breakthroughs and market expansion and has important research significance and application value.

The detection of malware is completely possible using a cloud distributed parallel environment for distributed parallel detection of malware. The maximum independent set algorithm is used to optimize the network structure, avoid network congestion, and balance the load. Distributed parallel detection of suspicious software through the nodes of the distributed network improves the efficiency of detection and finally forms an analysis report that is sent to the client. This paper summarizes the malware detection methods based on machine learning and proposes a framework for malware detection by machine learning guided by distributed system technology to address the problems of slow response and lack of global analysis due to the existing malware detection methods that are difficult to cope with the large-scale network growth. The main contributions are as follows: (1) A new framework for distributed malware detection in a large-scale network environment is proposed. The newly emerged malware is analyzed, and its features are extracted and then updated to all nodes of the system, thus quickly achieving the detection of this malware in the whole system jurisdiction. (2) A new machine learning malware detection method is proposed to reduce the dimensional of the malware feature vector using the PCA algorithm. The random forest algorithm is further designed to obtain detection results by classifying malware using its robustness and strong fitting ability. (3) Experiments are conducted on Ember 2017 and Ember 2018

datasets, and the outstanding performance of the proposed method and its detection capability in large-scale networks are verified by comparing and analyzing the experimental results. Finally, a conclusion and future perspectives are provided.

2. Related Work

2.1. Distributed Architecture. Earlier detection of computer malware was done on the host computer in a completely isolated and controlled environment that did not require collaboration. However, network malware spreads throughout the Internet, an environment that cannot be isolated and is not fully controllable. Especially today, with the rapid growth of the Internet and the increasing size of networks, it is impractical to isolate each network in turn and then remove the malware. Collaboration can also include analysis of anomalies, notification of response methods, and feature broadcasting. Qiu et al. [5] gave the main problems and countermeasures to be faced by distributed mining of data streams, which is one of the earlier and more comprehensive pieces of literature on distributed mining techniques for data streams. One of their important points is that effective use of limited computer processing resources to solve the knowledge discovery problem of potentially infinite data requires finding a distributed solution that balances cost and accuracy, while Aslan and Samet [6] illustrated that distributed mining of data streams requires a comprehensive consideration of distributed computing, memory buffering, and node interaction costs from the perspective of performance optimization. They also designed a hierarchical mining architecture for data streams; that is, the mining system consists of several local nodes and one central node, the primary patterns are formed by parallel mining of local nodes, and the central node then generates global knowledge patterns. These studies, especially the mining idea of cost-accuracy balance, are also the key issues to be considered when designing the mining model in this paper. Of course, the distributed mining of data streams is concerned mainly with how to use distributed parallelism to solve the knowledge discovery problem of large-capacity (single) data streams. In fact, with the popularity and growth of network-based computer applications, multiple data streams with multiple nodes independently clustered but logically related have become another important data form, that is, distributed data streams. Admittedly, most of the current literature on distributed data stream mining is still focused on the study of key scientific problems and solutions, but as the scientific problems become clearer, some related mining architectures and methods have been discussed in recent years. Alazab et al. [7, 8] proposed a mining algorithm DS-means for distributed data streams. The typical unsupervised learning algorithm K-means is used in both local and global clustering of DS-means. However, in order to accommodate the capacity fluctuations of different data streams on different nodes and different time periods, the number of global class clusters is dynamically adjusted based on the clustering results of local nodes before global node clustering. Liu et al. [9] constructed a distributed

multinode mining architecture for the problem of mining frequent itemsets for distributed data streams and designed corresponding algorithms to implement the key operations of the architecture. The work in this paper draws on these distributed data stream mining architectures and methods. In short, distributed data stream mining is a natural extension and development of the research on data mining and its data stream mining techniques, and the current research is mainly focused on algorithm design. However, in the face of the growing demand for big data, systematic research on the mining architecture of big data and its core mining models and algorithms under the related architecture becomes a problem that must be faced. In this sense, distributed data streams can provide an ideal model of data organization for big data in many application environments, and mining algorithms for distributed data streams can be used as one of the core techniques to address big data mining.

In fact, as the concept of big data is getting hotter and hotter, some scholars or research institutions have started to explore the knowledge mining problem in big data experimentally. Gupta and Rani [10] argued that the mining of big data needs to study different algorithms at different processing stages and designed corresponding algorithms for some key steps. Naeem et al. [11] discussed the problem of big data analysis from the data mining viewpoint and explored the problem of big data based on data-driven strategies. And based on the data-driven strategy, they explored the core problems in big data mining and designed an architecture and core model for big data mining. Pan et al. [12] argued that big data has two data forms, batch and streaming processing, and analyzed the key technologies in big data streaming computing in a more systematic way. Big data mining is also a balanced optimization problem of cost and accuracy, where improving the accuracy of distributed mining with reasonable communication cost is one of the key scientific problems. Euh et al. [13] proposed an idea of mining microcluster patterns in data streams, that is, extracting statistical values such as points and means of each cluster after performing clustering algorithms on a data stream to form the so-called microcluster patterns. Amer and Zelinka [14] designed the statistical information of microcluster-like data into a tree structure that grows with time to maintain it. Both of these works are oriented to single data stream mining. A related approach directly oriented to distributed data stream mining comes from the work of Omer et al. [15]. Although microcluster mining is effective at local nodes, microcluster patterns are not suitable as global patterns because they lack sufficient predictive power, that is, the ability to classify unknown data. In fact, in distributed data stream classification mining, although data are collected and decentralized at local nodes, the data streams scattered at local nodes are interconnected, so discovering global classifiers for multinode sharing and prediction is one of the main tasks of classification mining for distributed, streaming big data. In addition, since distributed mining has to go through local pattern mining and data transmission before global pattern mining, the multistep processing may lead to degradation of data quality, so the selection and design of

global patterns must also consider the noise immunity performance. Based on the above two points, considering the high predictive capability and better robust performance of the integrated learning technique, this paper will study the malware detection methods suitable for the distributed approach by drawing on the existing integrated learning technique.

2.2. Based on Machine Learning Methods. Early detection and analysis of malware were performed by security experts using various tools to assist in tracking, tuning with disassembly to obtain the behavior of the software, and then combining their experience to give analysis results. This manual analysis method is technically mature, and if security personnel are experienced, the possibility of error is very small. But the analysis efficiency is very low and has been unable to meet the rapidly growing needs of malware analysis, so security vendors and researchers have proposed a variety of automatic detection and analysis methods. Now the research on malware is divided into four main categories: (1) malware detection, the study of the real function and purpose of software execution; (2) analysis, according to the detection results of the detection object automatically classified; (3) prevention, to prevent malware running in the computer; (4) recovery, if the malware has invaded the system, remove it from the system and repair the damaged system; the ideal antimalware system should include all of the above functions. However, the current security software for ordinary users only has the functions of categories (1) and (3), and only professional security software has the functions of category (4), while the software of category (2) is still in the research stage. According to the different perspectives, malware detection methods are divided into two types of static detection and dynamic detection; the following are the two types of methods, the existence of the differences and their respective advantages and disadvantages.

Static detection is a method that gives analysis results based on the characteristics of the program being tracked and analyzed without being run. The earliest use of static detection technology is currently used by ordinary users of various types of antivirus software, which has a history of 30 years. It is established manually by the security vendor to create a malware feature library, the antivirus software only needs to compare the software being detected with the software in the malware feature library, and you can conclude whether it is malware. The sent kind of eigenvalue scanning technology is very mature, and the false positive rate is very low. However, the establishment of a malware repository requires security vendors to invest a lot of manpower and resources and keep it updated in real time. And with more and more malware, the malware library is becoming more and more huge and large, the scanning and analysis speed is getting slower and slower; more seriously, this technology can only identify malware captured by the vendor and the timeliness is relatively poor, so it is no longer the focus of industry research. Currently, widely studied static detection technology is

the use of machine learning algorithms for detection and classification.

Cai et al. [16] first introduced Bloodhound technology, which uses heuristic algorithms to analyze software, with a certain degree of machine learning, can use existing knowledge of the software, and has not been manually analyzed to give analysis results; then the eigenvalue scanning method is more time-sensitive, no longer completely dependent on malware libraries. Kouliaridis et al. [17] proposed a standard compressed distance method to measure the similarity between malware and then classify them. Hei et al. [18] used the system API call sequence as the basis for determining malware, which can be able to restore the deformed call software dynamically, partially solving the software deformation problem. Liu et al. [19] used the program data stored in the PE format file to analyze and obtain the DLL file from which the program calls the system. Huang et al. [20] proposed an inference tree-based analysis method. They used the k-testable tree to differentiate the data flow of software to obtain a relationship graph between the data and use this as a basis for classifying malware. Liu et al. [21] studied the use of variable-length n-grams as classification features, and the experimental results showed that variable-length n-grams as classification features can achieve higher accuracy. Ozsoy et al. [22] studied text classification techniques to improve the traditional feature code scanning method and proposed an automatic n-gram-based detection method, which has leading experimental results. Galal et al. [23] conducted an experimental comparison of malware analysis techniques used to test a variety of mainstream classifiers, including KNN and NB. Naval et al. [24] proposed an attribute-based automatic detection method. It first statically extracts the n-gram information of the program and then uses the information gain value as the criterion for feature selection; then, the extracted features are reduced to eliminate redundant features, and the optimized feature vector is fed into the classifier to achieve automatic detection of malware. The static detection method based on an integrated neural network is also proposed, and the feature selection technique is introduced in the construction of the neural network to combine the perturbed input attributes and the chemistry training data to generate the neural network. The conclusion obtained is that the integrated classifier is better than the single classifier. Although scholars at home and abroad have made unremitting efforts in the field of static detection and achieved many results, static detection methods have never achieved practical results. In addition to the design of classification algorithms, there is another difficulty that can never be solved: the data used for classification cannot be guaranteed to be real.

Dynamic detection and analysis are a method of letting the program under test run and then extracting features for detection and analysis. The features here mainly refer to the behavior of the program or the result of the behavior, so it is also called behavioral feature detection and analysis. The study of dynamic detection and analysis is actually also divided into two parts: one is the acquisition of behavior (i.e., detection), and the other is the analysis of behavior. The acquisition of behavior is also divided into two types of

technical solutions. Another type of real-time monitoring method is the smart Hook technology inserted between the program and the operating system or hardware device, which can directly monitor all the persistent behavior of the program, which is also the mainstream monitoring method. Using Hook technology, the industry has developed a number of types of widely used real-time behavior monitoring software; they have different functions, and each has its own strengths and becomes an indispensable helper for researchers to study the behavior of viruses. Kouliaridis et al. [17] proposed a new method of scanning memory areas using hardware. Huang et al. [20] proposed a technique to check malware based on component integrity, whose kernal also uses digital signature techniques and downloaded software components from the web. Das et al. [25] proposed a method to monitor malware using hardware virtualization and software virtualization techniques. Idrees et al. [26] proposed a new approach to protect against malware injection by using hardware secure return address heap-finding techniques to achieve this dedicated hardware heap-finding to protect legitimate programs at the hardware level at a minimal cost to the system, complementing traditional software buffer overflow prevention methods. To analyze the kernel program, Omer et al. [15] proposed to compare whether the system kernel is modified by malware by taking a snapshot of the kernel environment. This method is not easily detectable by malware. To detect malware, Euh et al. [13] proposed a technique called confusion call detection. The algorithm was improved by Idrees et al. [26] by using a "sliding window" to improve the efficiency of the algorithm. The basic idea is to measure the similarity between sequences by calculating edit distance and then use the KNN algorithm to cluster the malware. Demontis et al. [27] used the bag-of-words model to parse program behavior, automatically extract the feature words, and convert these feature words into vector space data and then used support vector machines to achieve the automatic classification of malware. Yewale and Singh [28] proposed the use of an expert system for malicious behavior determination, which evaluates all behaviors as a whole, considers their potential maliciousness, and calculates the overall result. However, they only proposed such an idea and did not give a concrete implementation of it. Based on these studies, experimental systems have been developed in several research institutions. In general, although the current program behavior-based pickup and analysis method is not as commonly applied as the static analysis method, it has some advantages that the latter does not have, so the development prospect is more promising, and it is also the main direction of researchers at present.

3. Method

The overall structure of our proposed approach is shown in Figure 2. In Section 3.1, we first describe the architecture, components, and main functions of the proposed distributed system. For the subnodes in the distributed system, we describe in detail their algorithms for performing feature extraction and random forest-based malware detection. In Section 3.2, we describe in detail the proposed random forest-based malware detection method.

3.1. Distributed Architecture.
The system uses a hierarchical model of distributed detection and collaborative analysis management. In the system, detection nodes are arranged on the nodes of each large-scale network, and the probes of the nodes are placed on the network entrances and exits to be monitored. The node module consists of a feature detection module, anomaly detection module, and analysis control module. The feature detection module detects various malware with known characteristics, and the anomaly detection module is responsible for finding anomalies. They are connected through the analysis control module, while each node communicates and collaborates with each other through the analysis control module, thus framing a distributed and collaborative malware detection system on a large-scale network. The following describes the model for constructing a distributed malware detection system based on node detection.

3.1.1. Node Detection.
At the access point of the network, the data is bypassed into the processing system through the port mapping of hub or switch. At each node, the analysis function of the node is divided into two parts, and the data captured by the probe is sent to both anomaly detection and feature detection modules: (1) processing of known malware, using the existing feature library to detect known malware in the network; (2) analysis of the data using the anomaly discovery module, from which malware not yet in the library is found. The results of the analysis and the data to be further analyzed are sent to the analysis console. As the functions of the two modules differ greatly, the detection of known malware is relatively fixed, while the detection of the unknown can be selectively added or removed by means of anomaly discovery according to the actual situation. The detection points can be dynamically added and reduced according to the actual situation when constructing the network topology model without changing the network settings or increasing the load on the network. This division makes the system composition flexible and extensibility.

3.1.2. Distributed Topology.
Due to the system's distributed detection approach, the system is highly extensible, and the architecture is extremely flexible, allowing the system to scale effortlessly when monitoring across networks, thus enabling distributed interoperability, which is of great importance for Internet-wide malware detection. The distributed collaboration of the system is shown in Figure 3.

In Figure 3, each node is able to perform the detection task independently; at the same time, it can communicate with other nodes through the collaboration channel. When a new malware appears in a subnet, it is captured by the detection nodes in the subnet and quickly analyzes the characteristics; then, it is notified to other nodes through the collaboration channel so that other nodes have the ability to detect such malware. The connection relationship between
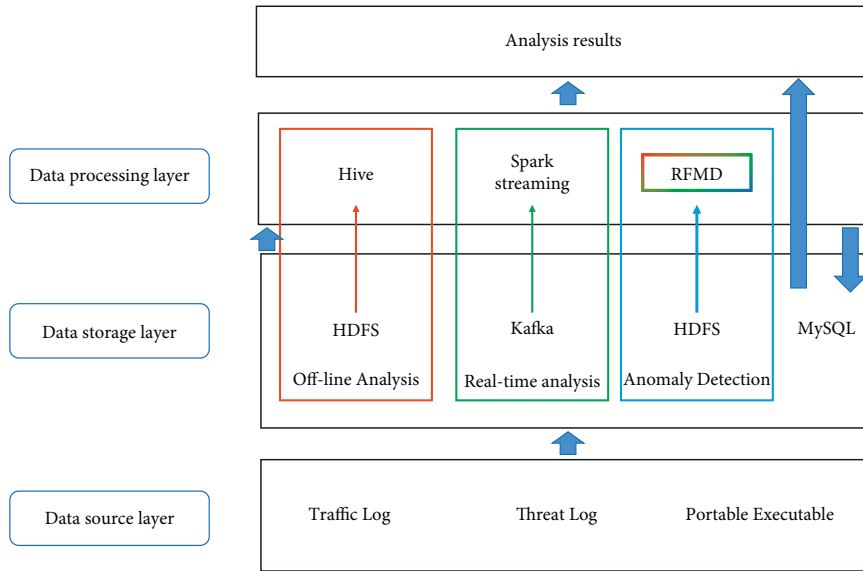
Figure 2: Distributed system technology-guided machine learning in malware detection architecture.
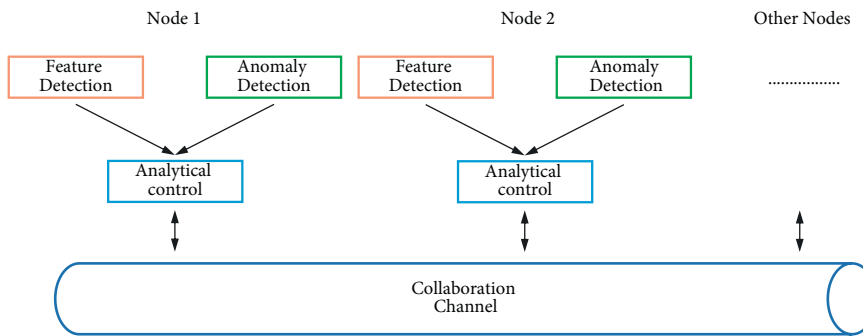
Figure 3: Sketch map of distributed cooperation.

analysis and control modules can be a simple peer-to-peer relationship, or a certain loosely coupled hierarchical relationship can be structured, which does not affect the working of the system. High-level analysis control modules are simply application-based definitions that function and implement exactly the same as other analysis control modules. The analysis results of multiple single nodes can be reported to the analysis control platform; the analysis console can synthesize the situation of each node for analysis, thus giving a global view of the analysis. In turn, the regional analysis console can send the data that needs to be analyzed to the advanced analysis console so as to obtain powerful analysis support and improve the system analysis performance when the analysis task cannot be completed. In each node of the backbone network, a packet-catching probe and malware analysis platform are placed to analyze the captured data, from which anomalies are discovered and the results are sent to the analysis control center in the region. Each regional node is connected to a regional analysis control center through a dedicated line; within the entire controlled network, all analysis platforms are connected to an advanced analysis control center.

3.1.3. System Functions. The functional components of the malware signature detection module and anomaly detection module, as well as the analysis and control module, are described as follows:

(1) Malware feature detection module: As there are many various known malware currently prevalent on the network, their behavioral characteristics are well known, and their signature codes have been obtained; their detection is performed through known modules, such as CIH. The block diagram of the known detection module is composed as shown in Figure 4. The data stream enters the detection engine, which analyzes the data and writes logs if known malware is found and notifies the response module to perform analysis of the logs to generate reports for publication on the Web. The functions of the modules are as follows.

(2) Detection engine: According to the content of the feature database, analyze the data obtained by the capture platform; if a match with known features is found in the data, notify the response module and carry out the corresponding response action; at the
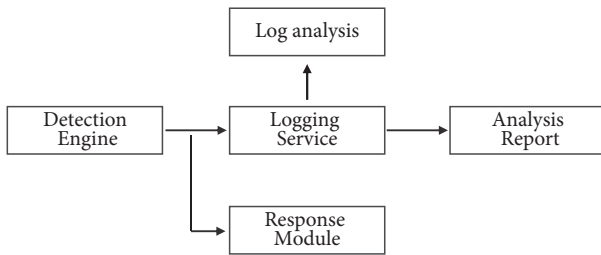
FIGURE 4: Signature detection of malware.

same time, record the generated logs. Log information includes source IP, destination IP, source port, destination port, and type of malicious code. Log service: accept the analysis results generated by the engine, write the results into the database, and publish them through the Web. Log analysis: analyze the log database, generate reports and statistics based on the data, and provide queries.

(3) Statistical methods are the simplest and most feasible way to detect anomalies. Specific statistical analysis methods are currently a hot research topic and are under rapid development. Statistical analysis methods first require the creation of a statistical description that summarizes a number of measured attributes (such as source address, destination address, and destination port counts) and records the values of these attributes during normal use. The width of the methodological confidence interval for the statistics depends on the user setting different confidence levels and does not require previous knowledge about anomalous activity. Statistical anomalies are found from the observed values, and the confidence intervals change themselves as more data are measured. Statistical analysis is based on the usual statistics, setting confidence intervals for the transmission on the network; if the measured attributes are not within the confidence interval, anomalous behavior is considered to have occurred. The advantage of statistics is that the algorithm and implementation are very simple and can detect unknown and more complex intrusions; the disadvantage is that there are false positives, high leakage rates, and low sensitivity.

(4) Sampling: Real-time detection of anomalous behavior for large-scale backbone networks is constrained by system performance, mainly (1) processing speed, including the speed of network adapter hardware, protocol stack reduction, and detection algorithms; (2) storage capacity, because the backbone network bandwidth is very high and the data per second is in GB if stored. In this case, the sampling detection method can be used. This detection does not require the use of traditional methods to measure and record all traffic data. Using sampled message measurements is sufficient for real-time needs rather than measuring all traffic information. Sampling can be done at regular intervals or

according to certain sampling methods. Sampling methods include random sampling and Poisson sampling. In this paper, we adopt distributed sampling and fixed-time source sampling. The sampling method is used to obtain anomalies in network traffic and network behavior. The data obtained by sampling is used for statistical analysis, for both anomaly discovery and more complex content analysis. The sampling parameters are modified so that the samples obtained by sampling describe the statistical properties of all traffic as well as possible. The advantage of sampling is that the samples can be processed and analyzed in a more focused and detailed manner; the disadvantage is that some anomalous information is easily missed.

3.1.4. Analysis Control Module.
The analysis control module needs to preanalyze the suspected data captured by the anomaly detection module in order to be able to discover the unknown malware in time, and if it is confirmed as malware, it has to perform sample analysis as soon as possible. The analysis object of analysis control is the suspicious data stream obtained from the anomaly detection module, and it is controlled by the next level of analysis control center, known detection module, logging service, and information distribution. The analysis control center centralizes all the control of the system, while all human operations are directly facing the analysis control center, and the analysis control center includes the work of all people. The work of each module of the analysis control center includes the following (see Figure 5).

The sequential treatment process is as follows: (1) preanalyzing, which preanalyze the suspected data; (2) sample analysis to obtain its characteristic codes and activity characteristics; (3) analysis management platform, to conduct comprehensive analysis and evaluation of the samples; (4) system maintenance, the feature code obtained by analyzing the unknown anomaly is added to the feature library of each known detection module; (5) security support, including issuing notifications to users, generating reports, and releasing special killing tools. The features of new malware acquired by the analysis control platform through the analysis of anomalies can be added to the feature base of known detection modules through the updated model of the feature base, thus enabling such malware to be detected as known. As the requirements for emergency response are getting higher and higher and the analysis and control capabilities are not balanced everywhere, it requires a coordination problem that can coordinate all human and material resources to focus on quick problem solving and emergency response, and then it can effectively reduce losses.

3.1.5. Mechanisms for Collaboration.
In order to be able to collaborate among the nodes and to synthesize the analysis for easy management, communication between the nodes in the system is required. The communication includes (1) analysis of unknown malware, source data, and analysis results, (2) control information, and (3) notification of
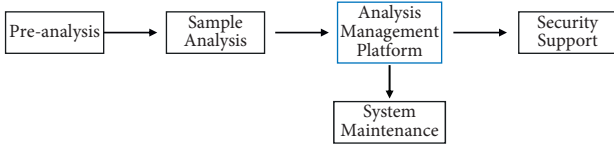
FIGURE 5: Model of anomaly detection.

FIGURE 6: Extensibility model of signature database.

analysis results and new response handling to each node. The most important feature library update is the communication between the nodes. As new malware appears faster and faster, the size of the feature library can expand dramatically and needs to be updated very timely. The system uses a centralized control structure that includes the console and its connected feature libraries for several known detection modules, as shown in Figure 6.

The console is responsible for updating all the feature bases, and two mechanisms of push and pull can be used when operating on the feature bases as follows: one update operation on the feature bases initiated by the console; the virus engine actively connects to the console to update the feature bases. When the analysis console analyzes the features of the newly discovered malware, it takes the initiative to update the feature base of the detection engine of each node, that is, push. Since new malware is not discovered in a short interval, pushing is not performed frequently, but to keep the feature base fresh, the higher-level analysis console is connected every longer period to obtain the malware features that need to be updated and remove the outdated feature codes. In this paper, we use a hierarchical and connected organization model. The analysis control center at the lower level communicates with and receives control from the higher level. At the same time, each analysis control center is autonomous, that is, fully qualified to work despite other failures.

3.2. Detection Algorithm.

In this paper, we propose a random forest-based malware detection algorithm (RFMD). By introducing machine learning techniques, RFMD is able to accurately detect new malware and its variants. The core modules include distributed feature extraction module, feature selection and transformation module, and Spark-based random forest classification module: (1) distributed feature extraction module: based on Spark parallel computing framework, this module performs batch analysis of PE files to obtain 56 features and form a 56-dimensional feature vector; (2) feature selection and transformation module: based on PCA algorithm to process the 56-dimensional feature vector and reduce the dimensional; (3) Spark-based random forest classification algorithm.

The file header information can also be used as program characteristics, but it is usually in portable executable (PE) format. The PE file header contains numerous information about the program under test, including the number of sections, section names, section sizes, timestamps, dynamic link library information, import table information, and
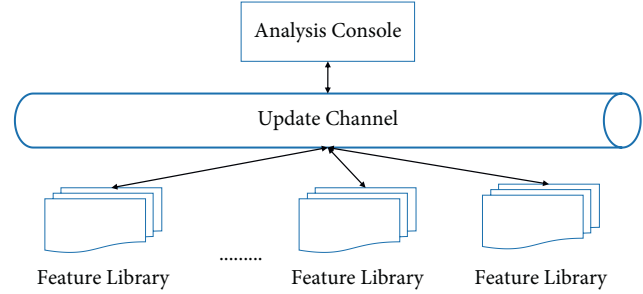
export table information. This information can be regarded as a macrodescription of the program to be tested. In this regard, the file header information is often used as an important feature for intelligent detection of malware. The core detection principle of the method proposed in this paper is that the feature identifiers contained in malware binary PE files can be used for automatic identification and classification of malware. The principle assumptions are described as follows: (1) In order to bypass the signature detection of AV, malware eliminates or circumvents the signature by variants, but the features contained in malware PE files are stable to a certain extent. (2) Malware causes its PE file features and structure differ significantly from benign software originating from commercial development environments. But this difference is only in functionality and does not affect its execution capability. Moreover, this difference can be used to distinguish malware from benign software to a certain extent. (3) Usually, malware uses shelling, encryption, and other obfuscation methods to bypass AV detection. (4) For rapid development as well as rapid mutation, low-level programming techniques usually trigger anomalies in PE file structure. (5) To implement functions commonly used by malware, some key combinations of API functions are usually information such as comments left in the software by malware authors, which are usually text strings. Based on the above six assumptions, the research and validation work in this paper consists of the following four steps, namely, potential effective feature design and extraction, feature simplification, feature contribution evaluation, and system evaluation. In order to enable the features in this paper to cover different segments of malware, features from different segments are extracted and subfolded, described as follows: information about the external structure of the file from the PE file header (DOS header and PE header) and information about the segments coming from inside the file structure are extracted. In summary, 56 potential features are extracted from the above description. Feature utility assessment, based on a training dataset, is performed in this paper to evaluate the ability of the 56 features in classifying malware and benign software samples, that is, classification contribution. The process results in a statistical model that describes the contribution of each feature to the classification and ultimately generates the metadata of the malware PE file. System evaluation: to validate the

effectiveness of the metadata of malware PE files for malware classification, based on the Spark distributed environment, this paper implements a prototype system, RFMD, to classify a collection of test PE files, resulting in malware and benign software based on 2 test datasets.

A random forest is a classifier consisting of multiple decision trees $\{h(x, \theta_k)\}$, where $\{\theta_k\}$ are mutually independent and identically distributed random vectors. The final class label of the input vector is determined by the combination of all decision trees. To construct k trees, we have to generate k random vectors $\theta_1, \theta_2, \ldots, \theta_k$, which are mutually independent and identically distributed. The random vectors θ_i can be constructed as a decision classification tree $h(X, \theta_i)$, which is simplified to $h(X)$. Given k classifiers $h_1(x), h_2(x), \ldots, h_k(x)$ and random vectors x and y, define the edge function:

$$mg(x, y) = av_k I(h_k(x) = y) - \max_{j \neq y} av_k I(h_k(x) = y), \quad (1)$$

where $I(\circ)$ is the demonstrative function. This marginal function portrays the extent to which the average number of votes for the correct classification of vectors x and y exceeds the average number of votes for any other class. The larger the margin, the higher the confidence level of the classification. Thus, the generalization error of the classifier is as follows:

$$PE^* = P_{x,y}(mg(x, y) < 0), \quad (2)$$

where the subscripts x and y represent that the error is in the X and Y spaces. Extending the above conclusion to a random forest, $h_k(X) = h(X, \theta_k)$. If the number of trees in the forest is large, the following theorem can be obtained using the law of large numbers and the structure of the trees. As the number of trees increases, for all random vectors $\theta_1, \theta_2, \ldots, \theta_k$, PE^* tends to

$$P_{x,y}\left(p_\theta(h(x, \theta) = y) - \max_{j \neq Y} p_\theta(h(x, \theta) = j) < 0\right). \quad (3)$$

It shows that the random forest does not overfit, which is an important feature of the random forest, and the generalization error PE^* will tend to an upper bound as the number of trees increases, which indicates that the random forest scales well to unknown instances.

Strength of classifier $h(X, \theta_i)$ is as follows:

$$s \leq E_{x,y} mr(x, y). \quad (4)$$

Assuming that $s \geq 0$, according to Chebyshev inequality, from equations (3) and (4), we can obtain

$$PE^* \leq \frac{\text{var}(mr)}{s^2}. \quad (5)$$

The upper bound on the generalization error of the random forest is derived and defined as

$$PE^* \leq \frac{\rho(1 - s^2)}{s^2}, \quad (6)$$

where ρ is the mean value of the correlation coefficient and s is the classification strength of the tree. From Theorem 2, the upper bound on the generalization error of the random forest can be derived based on two parameters: the classification accuracy of each decision tree in the forest, that is, the tree strength s, and the degree of interdependence between these trees ρ. When the correlation degree ρ of each classifier in the random forest increases, the upper bound on the generalization error PE^* increases; when the classification strength of each classifier increases, the upper bound on the generalization error PE^* increases. A correct understanding of the interplay between these two is the basis for our understanding of how random forests work.

The random forest works very differently from other classifiers because the model is trained with a set of random vectors $\theta_1, \theta_2, \ldots, \theta_k$. The design of this set of random vectors is to maximize the discretionary of the random forest to reduce the generalization error and improve its generalization ability. In the classification stage, the class labels are obtained by combining the classification results of all decision trees. The most currently used methods are voting and probabilistic averaging. For the test sample x, the predicted class label c_p can be obtained:

$$c_p = \arg\max_c \left(\frac{1}{N} \sum_{i=1}^{N} I\left(\frac{n_{h_i,c}}{n_{h_i}}\right)\right), \quad (7)$$

$$c_p = \arg\max_c \left(\frac{1}{N} \sum_{i=1}^{N} \left(w_i \frac{n_{h_i,c}}{n_{h_i}}\right)\right). \quad (8)$$

Equation (7) is the result of voting, and equation (8) is the result of probability averaging. N is the number of decision trees in the forest, and $I(*)$ is the schematic function. $n_{h_i,c}$ is the classification result of tree h_i for class C, n_{h_i} is the number of leaf nodes of tree h_i, and w_i is the weight of I decision trees in the forest. In this paper, we use voting to determine the class labels; that is, the test set goes over one side of each tree in the forest, the detection results of each tree for each software are recorded, and the final class labels of the software are those classes that receive more votes than some threshold.

4. Experimentation and Evaluation

4.1. Datasets. To effectively validate the proposed approach, we experimented with the publicly released Ember datasets from the security company Endgame, hoping for the application of distributed systems and machine learning techniques in malware detection. The data includes include two parts, including Ember 2017 and Ember 2018, and Table 1 shows the number of samples and categories of the datasets. According to the dataset's presentation, the collected malware samples were flagged as malicious by multiple detection engines, while benign samples were not flagged as malicious up to the time of collection, and this section uses these datasets as experimental data. An additional 7994 malicious samples and 7158 benign samples were

Table 1: Ember datasets statistics.

Datasets	Quantity (10,000)		
	Malignant samples	Benign samples	Unknown sample
Ember 2017	40	40	30
Ember 2018	40	40	20

collected during 2019 to validate and compare several major approaches.

4.2. Experimental Setup and Evaluation Metrics.
Experiments were conducted using the Python language and Spark distributed system, and the relevant environment configurations and software versions are listed in Table 2.

To evaluate the detection accuracy of the system, the performance of the model is evaluated using Accuracy, Precision, Recall, and $F1$-value ($F1\_score$). Precision denotes the proportion of the data judged as malicious code that is truly malicious, defined as Precision $= TP/(TP + FP)$. Accuracy denotes the overall precision of the detection system, defined as

$$Accuracy = \frac{TP + TN}{TP + TP + FP + FN} * 100\%. \qquad (9)$$

Recall, which indicates the ratio of the amount of relevant information retrieved from the database to the total amount, is defined as Recall $= TP/(TP + FN)$. $F1$-score, which is used to synthetically evaluate Precision and Recall, is the weighted summed average of Precision (P) and Recall (R). It is defined as $F1 = (2PR/P + R)$.

4.3. Detection Performance Comparison.
We have selected competitive approaches in machine learning-based malware detection for comparison. We conducted an analysis to validate several commonly used machine learning methods, such as the proposed RFMD, LightGBM, SVM, and K-means.

The models were cross-checked on the Ember 2017 test set, the Ember 2018 test set, and the 2019 sample set. Figure 7 shows the results of the performance evaluation, calculating Accuracy, Precision, Recall, and $F1\_score$, respectively. Based on the overall comparison, the following conclusions are drawn: (1) All models outperform their performance on the Ember 2017 test set than their performance on the Ember 2018 test set. According to the publisher's (Endgame) statement, the detection difficulty was purposefully increased when collecting the 2018 malware samples, thus producing this comparison in this experiment and reflecting that machine learning-based detection methods are still affected by the detection difficulty of the malware itself. (2) Comparing Accuracy, Precision, Recall, and $F1\_score$, we can see that "rf2017" outperforms "lgbm2018" on the Ember 2017 test set; on the Ember 2018 test set, the result is the opposite. This result indicates that the coverage of the training data has an impact on the model performance, and the detection ability of the model relies on the learning of training data so that the training data covers different types of samples as comprehensively as possible, which should

Table 2: Experimental environment.

Name	Version number or configuration information
Python	3.8
Java	1.8.0
Scala	2.12.10
Spark	3.0.1
Hadoop	3.3.0
Operating system	Windows 10
Number of cluster nodes	1 master node, 3 slave nodes
Node configuration	8 GB RAM, 4-core processor
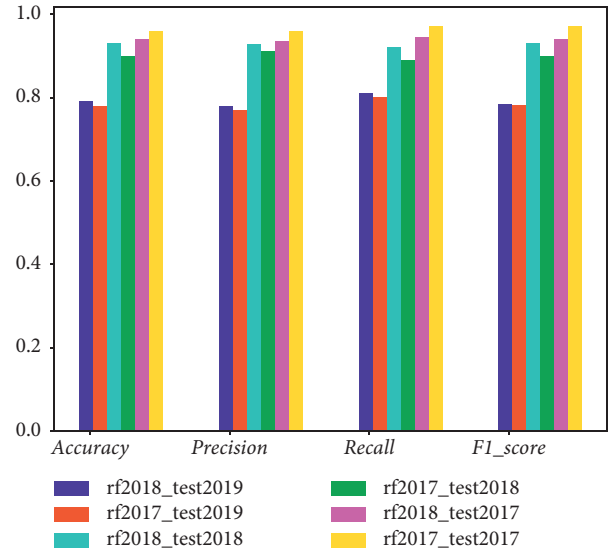
Figure 7: Evaluation metrics for RFMD models.

enhance the generalization ability of the model. (3) The average detection accuracy of all models on the 2019 sample set is about 80%, which is about 20% lower compared to the accuracy on the 2017 and 2018 test sets, indicating that the detection models are not sufficient to cope with the evolution of samples (including malicious and benign samples), such as malicious feature enhancement and adversarial detection. Therefore, the postmaintenance of the model is important, and the training set should be updated regularly; on the other hand, more robust detection methods are to be further investigated.

The proposed method, LightGBM, SVM, and K-means were used experimentally on the Ember 2017 datasets, and a comparison of the accuracy of the various algorithm models is shown in Figure 8. The experimental results also show that the proposed method performs best in Accuracy metrics.

4.4. Distributed System Performance Verification.
To verify the performance of the improved RFMD parallel algorithm, it was compared with the conventional RFMD algorithm running in a common stand-alone environment. The clustering results of the two algorithms are shown in Table 3. From Table 3, it can be seen that the number of samples correctly classified by the parallel algorithm is larger than
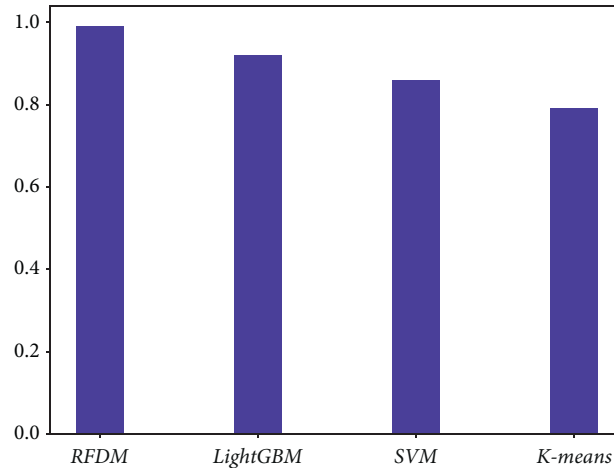
FIGURE 8: Accuracy comparison of different algorithms.

TABLE 3: Comparison of detection results.

	Number of samples (10,000)	Single model (%)	Distributed model (%)
Ember 2017	40	95.6	99.1
Ember 2018	40	91.4	98.6
Datasets	40	88.6	97.1

that of the traditional algorithm when different classes of Iris in the dataset are classified by the two algorithms; in the clustering of Ember 2017, the detection results of the two algorithms are 95.6% and 99.1%; in the clustering of Ember 2018, the traditional algorithm detects 91.4%, and the parallel algorithm detects 98.6% in accuracy; in the 2019's database, the traditional accuracy is 88.6%, and the parallel algorithm accuracy is 97.1%. The above results show that, compared with the traditional algorithm, the algorithm designed in this paper based on the distributed framework implementation has higher accuracy and better overall detection effect.

5. Conclusion

With the increasing growth and spread of malware, the traditional single-model approach no longer meets the needs of malware detection in large-scale network environments. For this purpose, this paper proposes a novel framework for malware detection by machine learning guided by distributed techniques, which overcomes the problem that traditional methods cannot adapt to large-scale networks. The proposed approach constructs a distributed framework that contains a data source layer, a data storage layer, and a data processing layer. The data source layer and data storage layer collect and store traffic logs, threat logs, and PE files, which are the basic work for detecting malware. In the data processing layer, we utilize Hive and Spark for offline computation and real-time analysis, respectively, while the proposed random forest-based malware detection algorithm for anomaly detection is implemented in a timely and effective manner. Experimental results on datasets Ember 2017 and Ember 2017 show that our proposed random forest-based detection method outperforms LightGBM, SVM, and

K-means. In distributed performance validation tests, our method significantly outperforms traditional single-model methods, and the real-time performance and accuracy of detection are effectively improved. The method provides technical support for large-scale malware detection on Web platforms. In the future, we plan to investigate deep learning-based malware detection and highly extendable distributed architectures.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] Y. Ye, T. Li, D. Adjeroh, and S. S. Iyengar, "A survey on malware detection using data mining techniques," *ACM Computing Surveys*, vol. 50, no. 3, pp. 1–40, 2017.

[2] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust intelligent malware detection using deep learning," *IEEE Access*, vol. 7, pp. 46717–46738, 2019.

[3] T. G. Kim, B. J. Kang, M. Rho, S. Sezer, and E. G. Im, "A multimodal deep learning method for android malware detection using various features," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 773–788, 2018.

[4] K. Grosse, N. Papernot, P. Manoharan, M. Backes, and P. McDaniel, "Adversarial examples for malware detection," in *Proceedings of the European Symposium on Research in*

Computer Security ESORICS 2017, pp. 62–79, Oslo, Norway, September 2017.

[5] J. Qiu, J. Zhang, W. Luo, L. Pan, S. Nepal, and Y. Xiang, "A survey of android malware detection with deep neural models," *ACM Computing Surveys*, vol. 53, no. 6, pp. 1–36, 2020.

[6] O. Aslan and R. Samet, "A comprehensive review on malware detection approaches," *IEEE Access*, vol. 8, pp. 6249–6271, 2020.

[7] M. Alazab, M. Alazab, A. Shalaginov, A. Mesleh, and A. Awajan, "Intelligent mobile malware detection using permission requests and API calls," *Future Generation Computer Systems*, vol. 107, pp. 509–521, 2020.

[8] J. Singh and J. Singh, "A survey on machine learning-based malware detection in executable files," *Journal of Systems Architecture*, vol. 112, Article ID 101861, 2021.

[9] K. Liu, S. Xu, G. Xu, M. Zhang, D. Sun, and H. Liu, "A review of android malware detection approaches based on machine learning," *IEEE Access*, vol. 8, pp. 124579–124607, 2020.

[10] D. Gupta and R. Rani, "Improving malware detection using big data and ensemble learning," *Computers & Electrical Engineering*, vol. 86, Article ID 106729, 2020.

[11] H. Naeem, F. Ullah, M. R. Naeem et al., "Malware detection in industrial internet of things based on hybrid image visualization and deep learning model," *Ad Hoc Networks*, vol. 105, Article ID 102154, 2020.

[12] Y. Pan, X. Ge, C. Fang, and Y. Fan, "A systematic literature review of android malware detection using static analysis," *IEEE Access*, vol. 8, pp. 116363–116379, 2020.

[13] S. Euh, H. Lee, D. Kim, and D. Hwang, "Comparative analysis of low-dimensional features and tree-based ensembles for malware detection systems," *IEEE Access*, vol. 8, pp. 76796–76808, 2020.

[14] E. Amer and I. Zelinka, "A dynamic windows malware detection and prediction method based on contextual understanding of api call sequence," *Computers & Security*, vol. 92, Article ID 101760, 2020.

[15] M. A. Omer, S. R. M. Zeebaree, M. A. M. Sadeeq et al., "Efficiency of malware detection in android system: a survey," *Asian Journal of Research in Computer Science*, vol. 7, no. 4, pp. 59–69, 2021.

[16] L. Cai, Y. Li, and Z. Xiong, "JOWMDroid: android malware detection based on feature weighting with joint optimization of weight-mapping and classifier parameters," *Computers & Security*, vol. 100, Article ID 102086, 2021.

[17] V. Kouliaridis, K. Barmpatsalou, G. Kambourakis, and S. Chen, "A survey on mobile malware detection techniques," *IEICE Transactions on Information and Systems*, vol. E103.D, no. 2, pp. 204–211, 2020.

[18] Y. Hei, R. Yang, H. Peng et al., "Hawk: rapid android malware detection through heterogeneous graph attention networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[19] Z. Liu, R. Wang, N. Japkowicz, D. Tang, W. Zhang, and J. Zhao, "Research on unsupervised feature learning for android malware detection based on restricted Boltzmann machines," *Future Generation Computer Systems*, vol. 120, pp. 91–108, 2021.

[20] X. Huang, L. Ma, W. Yang, and Y. Zhong, "A method for windows malware detection based on deep learning," *Journal of Signal Processing Systems*, vol. 93, no. 2, pp. 265–273, 2021.

[21] X. Liu, Y. Lin, H. Li, and J. Zhang, "A novel method for malware detection on ML-based visualization technique," *Computers & Security*, vol. 89, Article ID 101682, 2020.

[22] M. Ozsoy, C. Donovick, I. Gorelik, N. Abu-Ghazaleh, and D. Ponomarev, "Malware-aware processors: a framework for efficient online malware detection," in *Proceedings of the 2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*, pp. 651–661, Burlingame, CA, USA, February 2015.

[23] H. S. Galal, Y. B. Mahdy, and M. A. Atiea, "Behavior-based features model for malware detection," *Journal of Computer Virology and Hacking Techniques*, vol. 12, no. 2, pp. 59–67, 2016.

[24] S. Naval, V. Laxmi, M. Rajarajan, M. S. Gaur, and M. Conti, "Employing program semantics for malware detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2591–2604, 2015.

[25] S. Das, Y. Liu, W. Zhang, and C. Mahinthan, "Semantics-based online malware detection: towards efficient real-time protection against malware," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 2, pp. 289–302, 2015.

[26] F. Idrees, M. Rajarajan, M. Conti, T. M. Chen, and Y. Rahulamathavan, "PIndroid: a novel android malware detection system using ensemble learning methods," *Computers & Security*, vol. 68, pp. 36–46, 2017.

[27] A. Demontis, M. Melis, B. Biggio et al., "Yes, machine learning can be more secure! a case study on android malware detection," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 711–724, 2017.

[28] A. Yewale and M. Singh, "Malware detection based on opcode frequency," in *Proceedings of the 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, pp. 646–649, Ramanathapuram, India, May 2016.

Hindawi

Research Article

Fuzzy Support Tensor Product Adaptive Image Classification for the Internet of Things

Zhongrong Shi [ID],[1,2] **Yun Ma** [ID],[3] **and Maosheng Fu**[1]

[1]*Faculty of Electronic and Information Engineering, West Anhui University, Lu'An, Anhui 237012, China*
[2]*Anhui Yongcheng Electronic and Mechanical Technology Co., Ltd.,, Lu'An, Anhui 237000, China*
[3]*Faculty of Electrical and Opto-Electronic Engineering, West Anhui University, Lu'An, Anhui 237012, China*

Correspondence should be addressed to Yun Ma; mayun@wxc.edu.cn

Computer vision is one of the hottest research directions in artificial intelligence at present, and its research goal is to give computers the ability to perceive and cognize their surroundings from a single image. Image recognition is an important research direction in the field of computer vision, which has important research significance and application value in industrial applications such as video surveillance, biometric identification, unmanned vehicles, human-computer interaction, and medical image recognition. In this article, we propose an end-to-end, pixel-to-pixel IoT-oriented fuzzy support tensor product adaptive image classification method. Considering the problem that traditional support tensor product classification methods are difficult to directly produce pixel-to-pixel classification results, the research is based on the idea of inverse convolution network design, which directly outputs dense pixel-by-pixel classification results for images to be classified of arbitrary size to achieve true end-to-end and pixel-to-pixel high-score image classification and improve the efficiency of support tensor product models for high-score image classification on a pixel-by-pixel basis. Moreover, considering that network supervised classification training using deep learning requires a large amount of labeled data as true values and obtaining a large number of labeled data sources is a difficult problem in the field of image classification, this article proposes using a large amount of unlabeled high-resolution remote sensing images for learning generic structured features through unsupervised to assist the labeled high-resolution remote sensing images for better-supervised feature extraction and classification training. By finding a balance between generic structural feature learning of images and differentiated feature learning related to the target class, the dependence of supervised classification on the number of labeled samples is reduced, and the network robustness of the support tensor product algorithm is improved under a small number of labeled training samples.

1. Introduction

Image recognition technology is an important research branch in the field of computer vision, which aims to identify various potential objects in images using computers to preprocess, extract features, analyze and understand them. Traditional image recognition models can be divided into two parts: excellent feature extraction methods are robust in various complex environments, while classifiers mainly consist of some shallow machine learning algorithms for predicting the classes to which the features obtained by the extractor belong. For example, in the field of face recognition, researchers can improve the accuracy of a large

number of face image samples by simply feeding them to a model, which can then be trained over some time by deep learning models by more than a dozen percentage points over traditional models [1]. Excellent feature extraction methods are robust in various complex environments, while classifiers mainly consist of some shallow machine learning algorithms for predicting the classes to which the features obtained by the extractor belong. Target detection, one of the most challenging research hotspots in computer vision, is also a fundamental technique for solving more complex and advanced vision tasks such as semantic segmentation, target tracking, image description, and scene recognition. Deep learning techniques, which have emerged in recent years, are

a powerful method for learning feature representations directly from data and have already brought breakthroughs in the field of target detection. Currently, the most popular image recognition techniques include image classification, target detection, target tracking, and semantic segmentation, and a huge amount of problems in computer vision can be solved perfectly using these four techniques, either directly or indirectly. These techniques assist computers in extracting, analyzing, and understanding useful information from a single or a series of images, and the trend is for deep learning models to gradually replace traditional latent recognition models. However, while various popular algorithms or models embody excellent performance, there are still some general problems and challenges, such as the insufficient number of samples, poor target viewability, and slow convergence of models to train.

With the rapid development of artificial intelligence techniques and computer performance, image recognition based on machine learning algorithms has rapidly expanded from targeted application scenarios to become a standard scientific tool and applied to the full range of natural science and technology applications. In this article, we propose a novel adaptive image classification method, the particle swarm optimized fuzzy support tensor machine [2]. The method firstly calculates the fuzzy affiliation of each sample by fuzzy affiliation function to reduce the influence of noise points on the classification results; secondly, it uses particle swarm algorithm to perform parameter search for fuzzy support tensor machine; although it is more concise and easy to operate compared with other commonly used parameter optimization algorithms such as genetic algorithm and least squares method, the particle swarm algorithm still has the disadvantage of easily falling into local optimum in this article[3]. To solve this problem, the particle swarm algorithm is improved: firstly, the inertia weights are introduced into the particle swarm algorithm, which decreases nonlinearly with the number of iterations to improve the algorithm's optimality seeking ability, and secondly, the simulated image recognition algorithm is used to make the particles in the particle swarm algorithm forcefully jump out of the local optimum trap with a certain probability. The improved particle swarm algorithm greatly improves the efficiency of the optimization search and overcomes the blindness of parameter selection in the traditional classification model.

2. Related Work

Unlike pixel-based classification methods, object-oriented classification methods use the segmented objects or regions of an image as the minimum unit of analysis to compensate for the lack of contextual or spatial relationships in pixel-based classification methods. Object-oriented classification methods usually use segmentation before classification and obtain the classification results of the image by extracting features from the segmented results and then training the classifier.

In the literature [4], a nonparametric Bayesian hierarchical model is proposed for high-resolution remote sensing

image classification using a combination of object-oriented oversegmentation and hierarchical Dirichlet process model (HDP) and Indian buffet process (IBP), which solves a series of problems such as the traditional probabilistic topic model and ignores spatial information, and the number of topics has to be predetermined; in the literature [5], a remote sensing image is proposed for object-oriented Markov region penalty method for remote sensing images, using mean shift algorithm for image segmentation and establishing weighted region neighborhood map based on region size and neighboring region connection strength, while using region size and neighborhood strength features as penalty terms to calculate potential functions, and using maximum posterior probability to iteratively update joint probability distribution and likelihood function to obtain the final semantic segmentation results. This approach can better weigh the interactions between neighbors and obtain more macroscopic texture features; in the literature [6], after superpixel segmentation using the simple linear iterative clustering (SLIC) method, visual features are extracted and trained on the mixture components generated by the Dirichlet process mixture model through a multiple conditional random field model to obtain intermediate labels corresponding to visual features in the new feature space. Using the intermediate labels, further pixel semantic analysis is performed to establish the connection between low-level features and high-level semantics based on the spatial relationships taking into account the objects; literature [7] uses the watershed algorithm to record the process of gradual merging of oversegmented regions by region similarity comparison after oversegmenting the image using a binary segmentation tree, on which the ascending trajectory between the leaf nodes to the root node is the region evolution process obtained computationally, and the search for salient region content is achieved by finding the maximum value through first-order derivation of the evolution value, which improves the previous complex tree construction model and thus achieves image classification and target recognition more simply and efficiently. In the literature [8], a region of interest detection algorithm based on the MFF (Multiscale Feature Fusion) algorithm is proposed, which detects the region of interest in remote sensing images accurately and quickly by performing grayscale saliency analysis based on multiscale spectral residuals and directional saliency analysis based on integer wavelet transform on remote sensing images. In the literature [9], by performing Saliency Analysis of Cooccurrence Histogram (SACH) based on cooccurrence histogram for high-resolution remote sensing images and using a saliency enhancement method based on moving K-means clustering, clear region boundaries are established for the region of interest, while improving the immunity of the algorithm to noise. To reduce the computational complexity of region of interest detection for remote sensing images, the literature [10] proposes to achieve fast and efficient region of interest detection by segmenting high-resolution remote sensing images into superpixels and generating superpixel-level saliency maps using structural tensor and background contrast, and finally by superpixel-to-pixel-based saliency analysis. To extract high-quality regions of interest with clear

boundaries and no background interference from remote sensing images, a GLSA (Global and Local Saliency Analysis) algorithm based on global and local saliency analysis is proposed in the literature [11] for extracting residential regions in high-resolution remote sensing images. In addition, for common features of interest in high-resolution remote sensing images, such as residential areas, airports, aircraft, and ships, a detection algorithm based on joint multi-image saliency (JMS) is proposed in the literature [12], which processes multiple multispectral remote sensing images with similar spatial structure and spectral details by jointly using the correlation information between this set of remote sensing images to simultaneously detect the features of interest in this set of multispectral remote sensing images. The literature [13] proposes a region of interest detection algorithm based on superpixel segmentation and statistical significance analysis, which detects the region of interest in remote sensing images accurately based on the final generated significance map by fusing the statistical significance feature map based on histogram statistics and the information significance feature map based on information entropy analysis. The literature [14] proposes that candidate regions containing feature targets can be predicted by supervised learning models constructed from various salient features. Then a discriminative dictionary learning classifier based on sparse coding representation can be applied to the target candidate regions to detect feature targets in the scene, which greatly reduces the computational cost of traditional search strategies. For the airport detection problem of panchromatic remote sensing images, the literature [15] proposes to use a graph-based visual saliency model to locate the salient regions in the scene and obtain a top-down saliency map by making full use of the geometric prior knowledge of the airport runway, and finally by combining these two saliency maps, to predict the location of the airport more accurately. In the literature [16], a two-layer visual saliency analysis model is proposed to extract candidate regions of airports and aircraft, and a bag-of-words model based on dense SIFT features and Hu moment features are used to characterize the invariant features of airports and aircraft, and finally, the airport and aircraft targets in remote sensing images are accurately detected by support tensor machines.

3. Fuzzy Support Tensor Machine Adaptive Image Classification for the Internet of Things

3.1. Fuzzy Support Tensor Machine Theory. Support tensor machine based on statistical learning theory relies on its excellent learning ability and powerful generalization ability to have better results in dealing with problems with a small amount of sample data and nonlinear relationship between samples. However, there are still shortcomings in the support tensor machine model. The support tensor machine works by determining the support tensor with a small number of training samples and then finding a classification surface that can divide the samples and then classify the test samples. If there are incorrect or biased samples in the training samples to determine the support tensor, the support tensor machine cannot exclude these samples because they all have the same reliability for the test samples, so it is easy to be misled by these incorrect or biased test samples and establish the wrong optimal hyperplane, resulting in a decrease in the classification accuracy of the model. The biggest difference between standard fuzzy support tensor machine and support tensor machine is that the former has one more dimension than the latter, i.e., fuzzy affiliation. Hence, the support tensor selected by the fuzzy support tensor machine is not equivalent to that selected by the support tensor machine. To address this problem, fuzzy support tensor machines introduce the concept of a fuzzy affiliation function [17]. Each training sample is assigned a corresponding fuzzy affiliation according to its influence on the prediction result, with smaller affiliations for incorrect or biased samples and larger affiliations for correct samples, by which the problem that traditional support tensor machines are easily misled by isolated points is solved and the noise immunity of the model is improved.

As the complexity of the research problem increased, the degree of state in the problem was vague and could not be accurately described by traditional exact mathematics, and fuzzy mathematics was created to have a reasonable description of the degree of state of certain factors in the problem. The discipline was introduced in the 1950s and was mainly used to study some fuzzy problems. With the rapid development of artificial intelligence technology, fuzzy mathematics has been combined with various intelligent algorithms and is widely used in various fields. A fuzzy set is a basic concept of fuzzy mathematics. In the traditional notion of set, for an individual u and a set A, the relationship between them is that u either belongs to A or does not belong to A. These two results cannot hold simultaneously [18]. The relationship between an individual and a set, if expressed by a mathematical expression, should be

$$C_{ij} = \sum_{i=1}^{n} x_i^2 \sigma_i + x_j^2 \sigma_j + x_k^2 \sigma_k. \tag{1}$$

The eigenfunction cannot explicitly define data as belonging to a certain state or not belonging to a certain state. In the 1960s, researchers used feature functions to express an increasing number of classical sets by representing each data in the set as a fuzzy number, such that the domain of values of fuzzy sets was extended from the set of integers {0,1} to the set of real numbers [0,1]. Since the value of the fuzzy affiliation reflects the training points, for a certain class of defined affiliations, and the parameter j is a measure of the extent to which the support tensor machine misclassifies the samples, combining the two becomes a measure of how correctly the support tensor machine classifies data with different affiliations. Each training sample is assigned a corresponding fuzzy affiliation according to its influence on the prediction result, with a smaller affiliation for incorrect or biased samples and a larger affiliation for correct samples. By this approach, the problem that traditional support tensor machines are easily misled by isolated points is solved,

and the noise immunity of the model is improved. Transferring the processed data as input to the prediction model, the process of finding the optimal hyperplane for the classification model can be expressed mathematically as a quadratic program if the data transferred is linearly divisible.

$$E_{all} = \sum_{i=0}^{R} \left(E_{contest} \times x + Ee \times x + \varepsilon free \times x^2 \times d_i^{(1/2)} \right) \qquad (2)$$
$$+ \sum_{i=0}^{K} \left(4E_e \times x + \varepsilon decay \times x^2 \times d_i^2 \right).$$

Minimizing α in the objective function yields the quadratic counterpart of the pairwise plan:

$$I(X;Y) = \frac{K(x,y) \cdot \sum_{i=1}^{n} X_i Y_i}{M(x)} \cdot \frac{x - \alpha}{\sigma}. \qquad (3)$$

The major difference between a standard fuzzy support tensor machine and a support tensor machine is that the former has one more dimension, i.e., fuzzy affiliation, than the latter, so the support tensor selected by fuzzy support tensor machine is not equivalent to that selected by support tensor machine. If the relationship between the factors in the problem to be solved is nonlinear and a kernel function is introduced in the solution process, then the classification problem can be expressed in mathematical form as follows:

$$H(x) = \varphi \sum_{y \in \gamma} \sum_{x \in \chi} [p(x,y) \cdot \ln p(x,y) + Ax + Cy] + \lambda. \qquad (4)$$

One way to represent the above diffusion tensor is to use the covariance matrix of the Gaussian distribution, which mainly describes the diffusion process of water molecules in the tissue. The statistical scatter of the two distributions is used. In this case, tensor comparisons are made by measuring the Kullback–Leibler distance between the probability distributions afterward. The symmetric version is called J scatter, which was proposed by Wang and Vemuri and used for tensor distance measurements. It is shown as follows:

$$M(x) = \varphi \sum_{x \in \chi} \left[p^2(x) \cdot \ln p(x) \right] + Ax. \qquad (5)$$

Although Support Tensor Machines (STM) solve the overfitting problem in traditional SVMs, the rank-weight tensor is weakly expressive and this translation leads to poorer classification accuracy. The rank-weight tensor of STM is generalized to Tucker decomposition and CP forms to obtain stronger model expressiveness. However, the CP rank-decomposition causes an exponential increase in the number of parameters in the Tucker form, which suffers from dimensional catastrophe.

3.2. Fuzzy Support Tensor Machine Adaptive Image Classification for the Internet of Things.
Classification algorithm design has been a hot topic in the field of machine learning, pattern recognition, and computer vision. One of the most representative and successful classification algorithms is the Support Vector Machine (SVM), which has been highly successful in pattern classification by minimizing the

Vapnik–Chervonenkis dimensional and structural risk. However, standard SVM models are based on vector inputs and cannot directly deal with matrices or higher-dimensional data structures, i.e., tensors, which are very common in real life. As in Figure 1, a grayscale image is a two-dimensional matrix with height and width, which is a second-order tensor, while a multispectrum has multiple spectral bands and is a third-order tensor. When high-dimensional data are fed into the SVM, a common approach is reshaping each sample into a vector. A tensor can be seen as an extension of a matrix, which in traditional signal research can be considered as an array of different dimensions depending on the object of study. However, when the training data sample size is relatively small concerning the dimensionality of the feature vectors, this can result in overfitting and lead to unsatisfactory classification performance.

The traditional image recognition model consists of two parts: feature extractor and classifier. Feature extraction methods can be classified into texture features, shape features, bag-of-words model, sparse coding, local coding, and Fisher vectors. These feature extraction extracts feature from the image and then a set of numbers or symbols are used to represent certain characteristics of the depicted object in the image and finally, these features are recognized with the help of other machine learning methods. The classical recognition methods (classifiers) are support vector machines, decision trees, adaptive enhancement, plain Bayes, and some heuristic arithmetic [19]. These classical feature representation methods have a common feature that they all require a very specialized knowledgeable researcher to carefully design the model; however, this feature makes the model deficient in two ways: firstly, the researcher needs to spend a lot of effort to design different features for different recognition tasks; secondly, the practical application requires repeated validation and parameter tuning of the model, which is very costly to optimize.

Wavelet theory is an extended version of Fourier transform theory in which a signal is decomposed into wavelets and projected onto a set of wavelet functions. This differs from the Fourier transform, which decomposes the signal into sine and cosine components. Wavelet transform theory is popular in image processing, where it decomposes the input image into a set of images with various resolutions while reducing redundancy in the image representation [20]. The parent signal components are decomposed in an extended signal variant or a shifted wavelet. Two basic properties must be satisfied for a wavelet to be considered a wavelet. The information flow used for a single-level or one-level 2D image decomposition scheme is illustrated in Figure 2.

The inverse 2D wavelet transform used to reconstruct the image involves column upsampling and filtering for each subimage using low-pass and high-pass filters. The initial source image is constructed using the low-pass filter L and the high-pass filter of the resulting image for row upsampling and filtering and the summation of all matrices. By examining the saliency type focusing on bottom-up, the images can be classified into two categories, spatial domain models and transform domain models, depending on
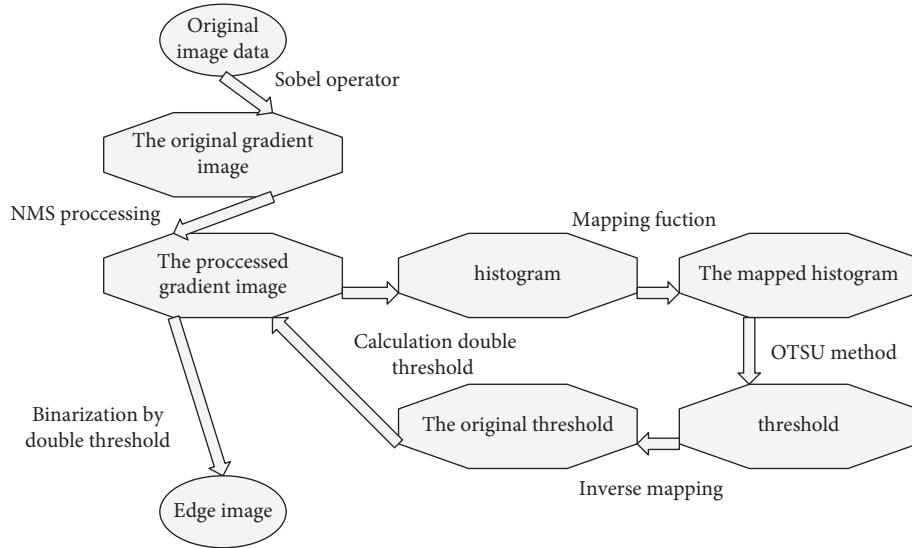
FIGURE 1: Picture and tensor representation.
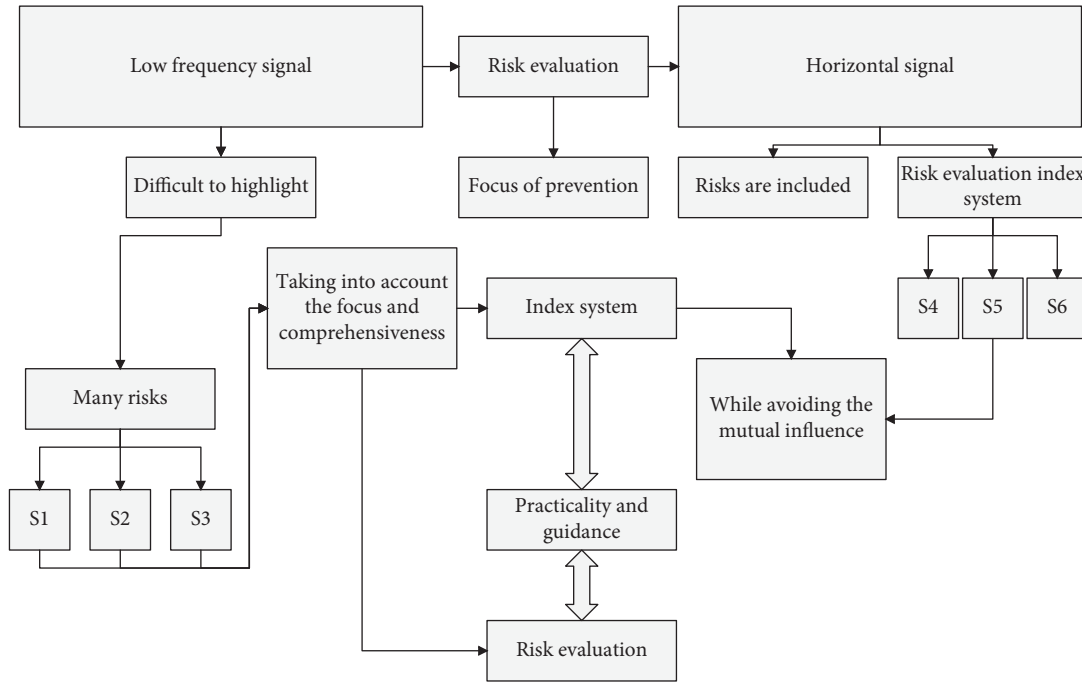
FIGURE 2: Wavelet decomposition of the two-dimensional image.

whether they are transformed in the frequency domain [21]. The so-called spatial domain saliency models process the image directly in the spatial domain and thus detect the salient targets or regions of interest in the image. Therefore, in the design of saliency models, regions in the image scene with unique color features or pattern features should have high saliency, while homogeneous regions in the scene should have low saliency. Features that frequently appear in the image scene should be suppressed. Salient pixels in an image should be clustered together rather than scattered throughout the image. Therefore, the Euclidean distance in spatial location of image blocks containing contextual

information about each pixel in the image is also important, because the distribution of image blocks in the background region is either far or near in space, while the distribution of image blocks in the salient target region tends to be clustered together in space. The saliency detection results can be further enhanced by incorporating the central prior knowledge of the salient regions. The saliency of pixel I in an image at a single scale can be defined as follows:

$$P_{i,j}{}^{k}(t) = \sum_{i=1,j=1}^{n,m} \left[t_{i,j}(k) \right]^{2} \left[\eta_{i,j}(t) \right]^{\beta}. \tag{6}$$

Furthermore, considering that image blocks in the background region are similar at multiple scales, in contrast, image blocks in the saliency region may be similar at only a few scales but not at all scales. Therefore, multiple scales can be used to further reduce the saliency of background pixels and enhance the contrast between salient and nonsalient regions. Unlike the spatial domain saliency model, the transform domain-based saliency model requires first transforming the image from the spatial domain to the frequency domain, then processing and analyzing the image in the frequency domain, and finally obtaining the final saliency detection results by transforming the analysis results in the frequency domain back to the spatial domain. For a two-dimensional signal like an image, by performing the Fourier transform on it, the resulting image amplitude spectrum clarifies the percentage of each sinusoidal component, while the phase spectrum of the image gives the position of each sinusoidal component in the graph. In the reconstruction of the image in the Fourier transform domain, the positions located in the horizontal or vertical directions with weak periodicity or homogeneity correspond to the positions of the candidate targets in the image, and thus it is known that the saliency information of the image is implicit in the phase spectrum of the image. Therefore, the saliency detection result of the image can be obtained by extracting the phase spectrum information of the image. The initial saliency analysis results are smoothed by a two-dimensional Gaussian filtering function $g(x, y)$ to obtain a visually superior final saliency map as follows:

$$e_j = -k \sum_{i-1}^{n} p_i \ln \frac{1}{p_i}. \tag{7}$$

From the above analysis, it can be seen that the PFT transform domain saliency detection model has the advantages of simple and easy algorithm and fast operation, thus giving fast saliency detection results for a given image, but the disadvantage is that the local saliency features of the image are not considered, and it lacks suitable biological psychological support and explanation. After each iteration step, the optimization progress of the solution needs to be measured by a predefined criterion that determines whether the current state is the best fit or not. Among the saliency analysis models for natural images, the center prior and the boundary prior are the two most widely used prior knowledge, which achieves good detection results in the saliency detection of natural images, thanks to the imaging mechanism of natural images, where the salient targets of natural images are usually at the location of the image center, while the boundaries of natural images usually do not have a distribution of salient targets.

Using superpixel segmentation methods, an image is segmented into many superpixels. Each superpixel contains a large number of spatially close neighboring samples that have a similar texture, color, luminance, and other characteristics. Compared to pixel-based hyperspectral image classification methods, the superpixel-based classification methods demonstrate good regional consistency.

The superpixel segmentation process is shown in Figure 3.

To alleviate pseudoboundaries that cause misclassification, we propose a new nonlocal decision-based region delineation method. In hyperspectral images, we usually consider that the samples in local regions belong to the same class, which is local information. However, nonlocal information is also very critical in hyperspectral images. This is because samples of the same class may also be located in different regions of the image. In nonlocal decision making, pixel pair similarity is extended to superpixel pair similarity, taking into account the structural information of the current samples. For those samples that are judged to be in heterogeneous regions by the local decision, the similarity between the current sample and the filter neighborhood samples is calculated. The current sample is represented by a global search for all similar samples. Then, this similarity is compared with the calculated adaptive threshold. If the similarity of all neighborhood pixels all is greater than the threshold, the current sample is judged to be in the homogeneous region and vice versa. This model includes two stages: the first stage is to enhance the input image, then input the residual network to carry on the supervised contrast learning, and get the pretraining model; the second stage is to fix the parameters of the pretraining model and the fuzzy support tensor machine is trained to get the prediction label. In the first stage, in order to enhance the discriminative ability of feature extraction, local information, nonlocal information, and generic structured features that come from unlabeled high-resolution images are also introduced, respectively. After an image of arbitrary size is input to the network as input data, the input data are convolved by each branch in the subnet separately utilizing dense convolution operations at different scales to associate the final extracted results, reduce the dimensionality through the transition layer, and then use the output data to (1) obtain the pixel-by-pixel classification results with the classifier to compare with the reference marker to calculate the loss; and(2) input them to the next subnetwork. The above process is repeated until the final objective function is obtained. The overall loss function is calculated jointly for all the objective functions and the network is trained by backpropagation of stochastic gradient descent. To facilitate training, the weights of the objective functions of all classifiers are learned in an alternating manner with the learning of other network parameters.

4. Experimental Verification and Conclusions

We compare the model in this article with Single Task Learning (STL), Tensor Train Multitasking (TT-MTL), and Tucker-based multitasking models. For a fair comparison, the same network architecture is used for all methods. In all experiments, this article sets the model format to $M = 3$ and $N = 2$. Since the model is harder to train when M, N is larger because of the presence of a fifth-order tensor. (Perhaps this can be solved by trying to decompose the larger cores further.) In this article, we use this relatively lightweight structure for our experiments. For the choice of rank, the
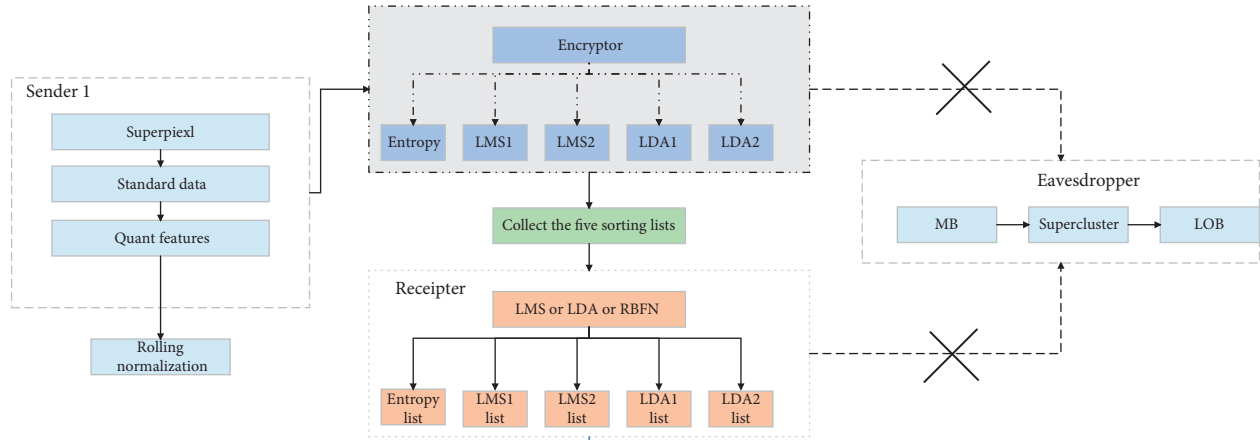
FIGURE 3: Superpixel segmentation process.

model parameters are extremely large when the rank is particularly large, so a relatively small rank (3, 4, 5) is used. The MNIST 10-class classification problem can be converted to a ten one-vs-all binary classification problem. This conversion allows the construction of a 10-task classification problem of the same kind, that is, by performing a softmax normalization on the ten classifiers before training. In this experiment, this article focuses on two performance metrics: one is the average accuracy of the ten binary classification problems, and the other is the accuracy of classifying a single digit by performing a softmax on the one-vs-all output of each task (multiclass classification accuracy). In this article, the first three convolutional layers are set to be hard-shared across all MTL models (for common feature extraction), and then the next FC layer is converted to a different multitask tensor network model format. In this article, we train with different sized subsets of the training dataset and test the model using the same test set (from 10% to 100% of the training set). As is shown in Figure 4, all MTL methods outperform STL for both more and less training data. Also, TT outperforms Tucker when the training data are small, while the results are reversed when the training data are large. The CTN proposed in this article outperforms all other methods.

In the whole experimental process, 100,000 training iterations are set in this article because the experimental images are few sample data, the choice of iteration parameters will affect the final model effect, and the accuracy and loss function of the model training can be analyzed to see in which interval the network reaches a stable equilibrium state so that the network training is in the optimal state, the accuracy and loss function curves are plotted and different network modules are designed to design the node embedding network, and the experimental results are shown in Figure 5. The comparison experiment of SE-ResNet structure and simple node embedding network access GNN can be obtained, the red line is the change curve of NSE-ResNet-EGNN network obtained by node embedding using SE-ResNet structure, and the blue is the change curve of simple node embedding network EGNN. This indicates that the extracted finite number of feature parameters can fully

reflect the transient response of the actual waveform pattern. From the accuracy change curve, we can see that the two approaches are almost the same in terms of the convergence speed in the early stage, but in terms of the final convergence result, the node embedding using the SE-ResNet network model can achieve higher accuracy, and in the subsequent iterations, the trend of the network curve is smoother compared to the simple node embedding network. From the loss change curve, we can see that the node embedding using the SE-ResNet network model makes the loss function fall faster in the training process and can maintain a smoother convergence effect compared with the original knot, and from the final convergence, the node embedding using SE-ResNet network model will have less loss, and we can see that by improving the dependency relationship between node channels, the performance of the node embedding network can be effectively enhanced by improving the dependencies between the node channels.

Though the previous experiment, it can be obtained that improving the node embedding network can increase the information encoded to the nodes, thus making the edge features describe the nodes more accurately. The network design of node update is also available in the GNN-Block module, so two sets of comparison experiments are set up in this experiment from the number of Conv-Blocks in the node update network and the network architecture, respectively. To verify the effect of the number of Conv-Blocks on the final results, we choose the node update framework as the comparison experiments, whose experimental results are shown as the yellow and red curves in Figure 6, respectively. From the accuracy change curve, we can get that increasing the Conv-Block of the node update network can improve the final classification accuracy within a certain range and keep the same convergence speed in the early stage, but in the final convergence result, the number of Conv-Block is proportional to the classification accuracy within a certain range using the node update network. From the loss variation curve, we can see that as the number of Conv-Blocks of the nodal update network increases, the loss function decreases faster due to the more parameters and better robustness of the network. Regarding network depth, and from the final
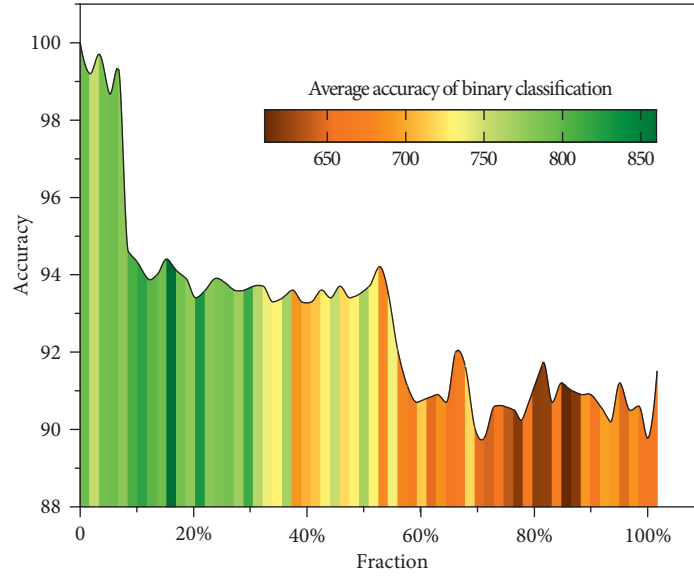
FIGURE 4: The average accuracy of binary classification for different algorithms with the same dataset.
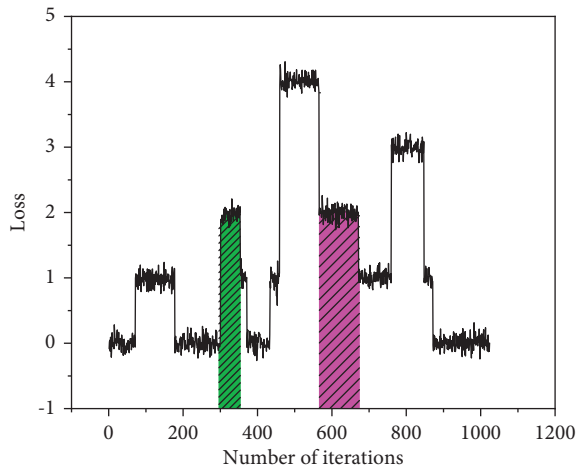
FIGURE 5: Classification accuracy and loss line of the model.

improvement results, increasing the complexity of the node update network can improve the few-sample classification performance within a certain range.

To visualize the effectiveness of the proposed algorithm, the distribution of representation coefficients and the corresponding normalized reconstruction residuals of the MFCARC algorithm are given. In our experiments, we selected the Indian Pines dataset for analysis, selected ten random samples from each feature class to construct the training dictionary, and randomly selected one sample from class 6 (grass/trees) of this dataset as the test book and then analyzed the representation coefficient distribution of this sample. Since the decomposition calculation process has a truncation process of redundant columns for the tensor factor matrix, some errors are inevitable while simplifying the calculation process. As shown in Figure 7, the distribution of the correlated adaptive representation coefficients based on different features exhibits the feature that the part of the representation coefficients with larger weights is

mainly concentrated in the category to which the test sample belongs. From the minimum representation residual criterion, it is known that the category to which the test sample belongs should have the smallest normalized residual value by comparing the reconstruction error of the test sample and each category of dictionaries. From Figure 7, it can be found that the correlated adaptive representation models based on spectral features, DMP features, and LBP features all make correct feature category determination for the test sample, but the correlated adaptive representation model based on Gabor features incorrectly determines the sample as category 4 (maize).

Most detection models are less effective in detecting smaller objects than in detecting larger objects. This is mainly because, after multiple layers of convolution, small objects may not retain any information in the feature mapping at the topmost layer of the model. Increasing the size of the model input (e.g., from 300×300 to 512×512) can help the model improve its performance in detecting small objects, with a 2.5 percentage point improvement in mAP for SSD and a 3.2 percentage point improvement in EAO. The analysis suggests that this is related to the multi-resolution detection layer proposed in this article, which gives different detection layers to detect objects of different sizes, such as the low-resolution detection layer used to improve the detection rate of small targets. The experimental results in Figure 8 are from seven smaller objects (bird, boat, chair, etc.) in PACAL_VOC2007, with XL, L, M, S, and XS in the horizontal coordinates denoting extra-large, large, medium, small, and very small, respectively, and the vertical coordinates denote the average detection accuracy of the model Here the different size pairs are produced by hand cropping postprocessing formation. From the figure, it can be seen that EA0 outperforms the base model SSD almost across the board and the advantage is more pronounced when the object size is of S and XS level. The analysis suggests that this is related to EA0's strategy of using different shapes
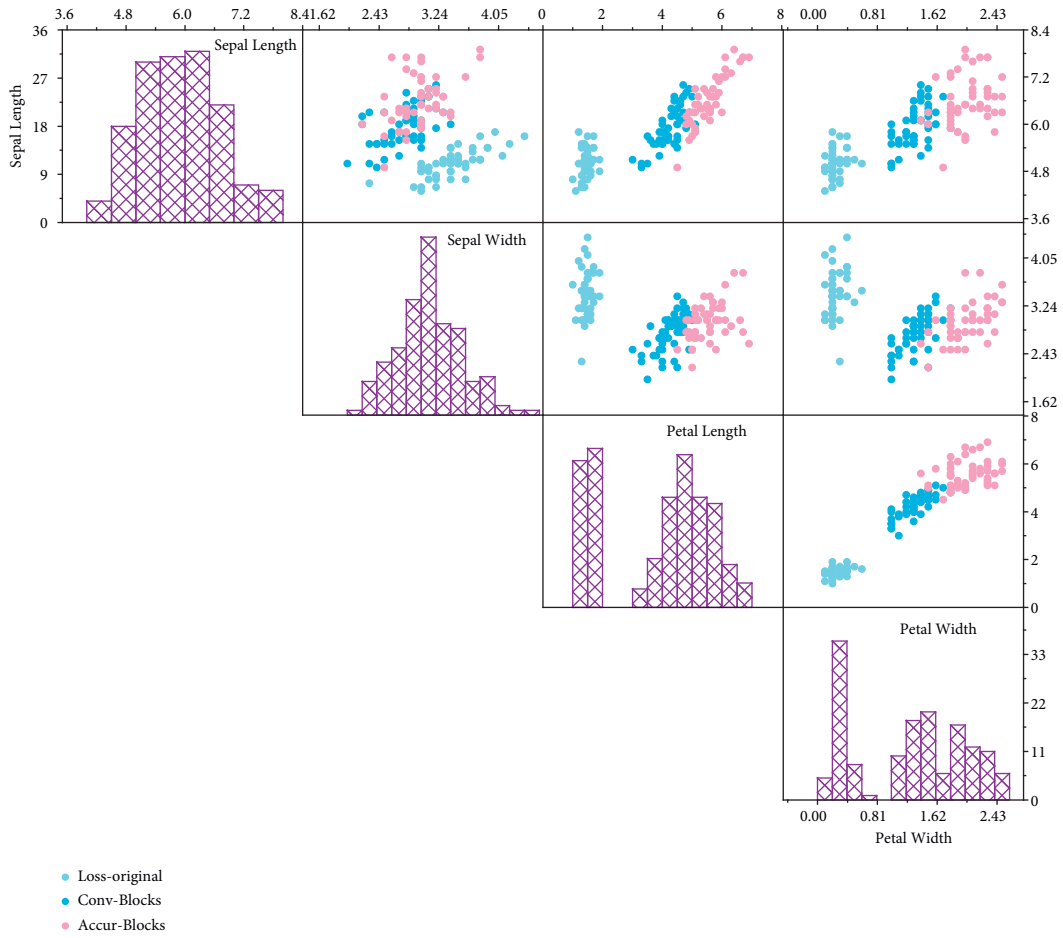
FIGURE 6: Classification accuracy and loss lines of the model under different GNN-blocks.
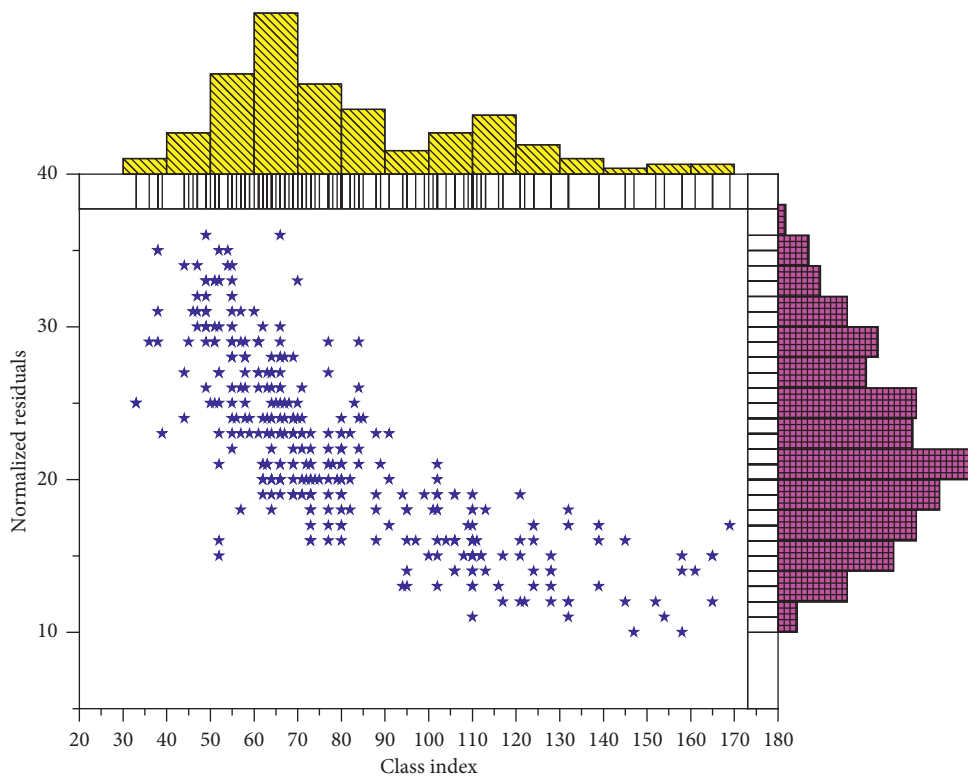
FIGURE 7: Normalized residuals for different algorithms.

Comparison of recognition accuracy
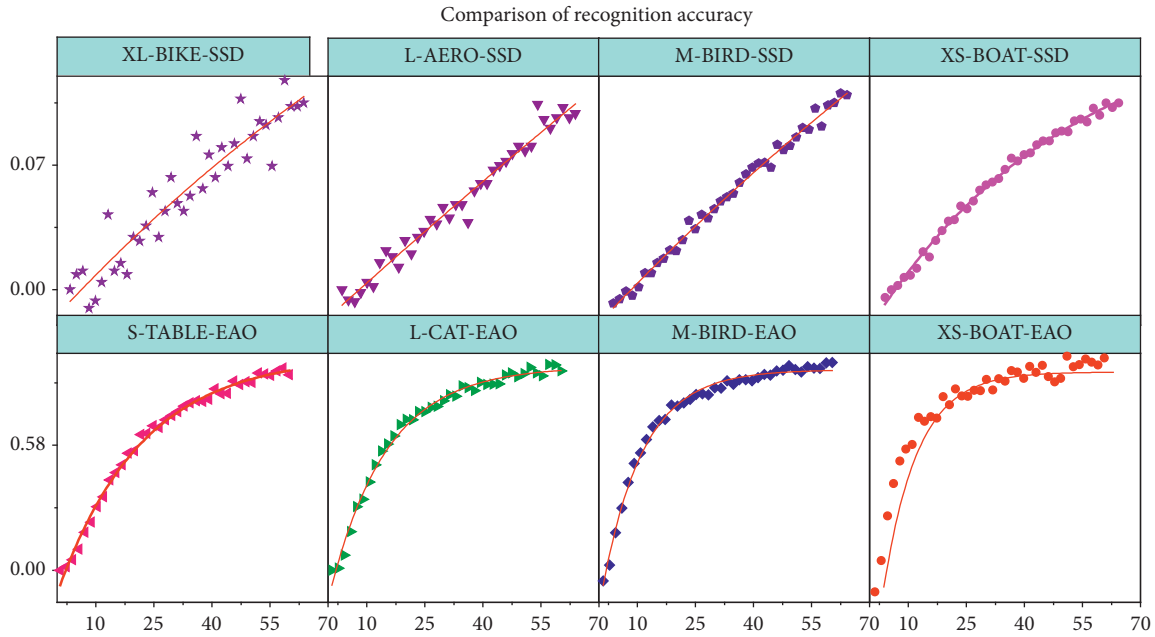
FIGURE 8: Comparison of recognition accuracy of different models for images of different sizes.

and numbers of a priori frames at different detection layers and assigning more detection frames to lower resolution layers based on the results of the cluster analysis. The best results achieved by EA0 in all scales of objects also reflect the lower sensitivity and greater robustness of the model to bounding box size than SSD.

5. Conclusion

Deep learning has been a great success in the fields of image recognition, speech recognition, and machine translation. Among them, support tensor machines have made breakthroughs in image detection, image classification, image segmentation, face recognition, video tracking, and other vision-related domains and have achieved great success in these fields. It is due to the powerful feature extraction capability of support tensor machines in image classification that more and more scholars are applying support vector machines to image classification. Tensor algorithms have different decompositions in several scientific fields. The different decomposition methods have their own advantages and areas of application. In this article, we propose an end-to-end, pixel-to-pixel IoT-oriented fuzzy support tensor product adaptive image classification method, introduce the background of the current topic of image classification and the significance of the research, as well as the current state of research on image classification, and draw out the difficulties faced by existing image classification methods by analyzing the characteristics of image classification and the advantages and disadvantages of existing image classification, providing strong realistic implications.

The accuracy of the prediction of the classification model established using fuzzy support vector machine and the selection of parameters of the algorithm have a great relationship, the more reasonable the selection of parameters,

the higher the accuracy. Therefore, in this article, to minimize the influence of parameter selection on the classification accuracy of IoT image recognition, nonlocal adaptive information is introduced, and the improved nonlocal information is combined with the fuzzy support vector machine to apply to the image classification research. This new approach to image classification is also the innovation of this paper. Experiments show that the model has better performance than standard RBMs in feature extraction and denoising tasks. Two visible and hidden layers of TRRBM are represented as matrix product states (MPS) and all computations can be done on a single kernel. This can significantly improve the computational complexity of the learning algorithm.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

This author declares that there are no potential conflicts of interest in this article.

Acknowledgments

References

[1] S. Smys, A. Basar, and H. Wang, "Hybrid intrusion detection system for internet of things (IoT)," *Journal of ISMAC*, vol. 2, no. 4, pp. 190–199, 2020.

[2] R. Rayhana, G. Xiao, and Z. Liu, "Internet of things empowered smart greenhouse farming," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 3, pp. 195–211, 2020.

[3] X. Wang and S. Gao, "A chaotic image encryption algorithm based on a counting system and the semi-tensor product," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 10301–10322, 2021.

[4] Y. Li, Y. Huang, M. Zhang, and L. Rajabion, "Service selection mechanisms in the Internet of Things (IoT): a systematic and comprehensive study," *Cluster Computing*, vol. 23, no. 2, pp. 1163–1183, 2020.

[5] J. Zhang and D. Tao, "Empowering things with intelligence: a survey of the progress, challenges, and opportunities in artificial intelligence of things," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7789–7817, 2020.

[6] M. Ruta, F. Scioscia, and G. Loseto, "Machine learning in the Internet of Things: a semantic-enhanced approach," *Semantic Web*, vol. 10, no. 1, pp. 183–204, 2019.

[7] M. Mafi, W. Izquierdo, M. Cabrerizo et al., "Survey on mixed impulse and Gaussian denoising filters," *IET Image Processing*, vol. 14, no. 16, pp. 4027–4038, 2020.

[8] C. Wang, "IoT anomaly detection method in intelligent manufacturing industry based on trusted evaluation," *The International Journal of Advanced Manufacturing Technology*, vol. 107, no. 3, pp. 993–1005, 2020.

[9] M. Yacin Sikkandar, B. A. Alrasheadi, N. B. Prakash, G. R. Hemalakshmi, A. Mohanarathinam, and K. Shankar, "Deep learning based an automated skin lesion segmentation and intelligent classification model," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3245–3255, 2021.

[10] D. Li, Z. Cai, L. Deng, and X. Yao, "IoT complex communication architecture for smart cities based on soft computing models," *Soft Computing*, vol. 23, no. 8, pp. 2799–2812, 2019.

[11] H. L. Nguyen, D. T. Vu, and J. J. Jung, "Knowledge graph fusion for smart systems: a Survey," *Information Fusion*, vol. 61, pp. 56–70, 2020.

[12] S. Sankaranarayanan, M. Prabhakar, S. Satish, P. Jain, A. Ramprasad, and A. Krishnan, "Flood prediction based on weather parameters using deep learning," *Journal of Water and Climate Change*, vol. 11, no. 4, pp. 1766–1783, 2020.

[13] Y. Sun, S. He, and F. Tong, "Media access control for narrowband internet of things: a survey," *Encyclopedia of Wireless Networks*, pp. 795–799, 2020.

[14] P. Wang, L. T. Yang, J. Li, J. Chen, and S. Hu, "Data fusion in cyber-physical-social systems: state-of-the-art and perspectives," *Information Fusion*, vol. 51, pp. 42–57, 2019.

[15] B. P. L. Lau, S. H. Marakkalage, Y. Zhou et al., "A survey of data fusion in smart city applications," *Information Fusion*, vol. 52, pp. 357–374, 2019.

[16] S. Vidyashree, "Smart shopping using android application," *Journal of Research Proceedings*, vol. 1, no. 2, pp. 243–252, 2021.

[17] R. F. Mansour, J. Escorcia-Gutierrez, M. Gamarra, D. Gupta, O. Castillo, and S. Kumar, "Unsupervised deep learning based variational autoencoder model for COVID-19 diagnosis and classification," *Pattern Recognition Letters*, vol. 151, pp. 267–274, 2021.

[18] Y. Han, C. J. Zhang, and L. Wang, "Industrial IoT for intelligent steelmaking with converter mouth flame spectrum information processed by deep learning," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2640–2650, 2019.

[19] S. H. Haji and A. B. Sallow, "IoT for smart environment monitoring based on Python: a review," *Asian Journal of Research in Computer Science*, pp. 57–70, 2021.

[20] R. Chandra Shit, "Crowd intelligence for sustainable futuristic intelligent transportation system: a review," *Iet Intelligent Transport Systems*, vol. 14, no. 6, pp. 480–494, 2020.

[21] G. A. Fotso Kamga, L. Bitjoka, T. Akram, A. Mengue Mbom, S. Rameez Naqvi, and Y. Bouroubi, "Advancements in satellite image classification: methodologies, techniques, approaches and applications," *International Journal of Remote Sensing*, vol. 42, no. 20, pp. 7662–7722, 2021.

Research Article

Network Fault Diagnosis of Embedded System Based on Topology Constraint and Data Mining

Tao Zhang (ID)

School of 3D Printing, Xinxiang University, Xinxiang 453000, Henan, China

Correspondence should be addressed to Tao Zhang; zhangtao@xxu.edu.cn

Maintaining the safe and efficient operation of network technology is an important development task of the computer industry. Topology constraint can optimize and combine the tracking results and select the target objects with better tracking performance to obtain the final tracking results and determine the target scale changes. Data mining technology can reduce the number of combinations to be detected, reduce the workload, and improve the timeliness and accuracy of the process of mining alarm association rules. Therefore, based on the summary and analysis of previous research results, this paper studied the network fault diagnosis of the embedded system method based on topology constraint and data mining. Firstly, a fault diagnosis topology model was established by constructing a topology search algorithm, which eliminated the filtering of association rules without topology relationship; the association rule-based data mining model was analyzed through the collection of network alarm data; the model algorithm was applied to the simulation experiment of network fault diagnosis of the embedded system and achieved good results. The results show that correcting rage of retrieval varies from 0.65 to 090 under different window sizes; the running time of the proposed method drops from 310 s to 35 s during 1–8 step/s of the sliding step, while the node degree ranges from 8 to 14 and diagnostic accuracy ranges from 0.97 to 0.94; the remaining alarm number increases from 0.5 to 3.5 threshold value, while the regular association number distributed in an interval of 40 to 140. The algorithm in this paper provides a reference for further research on network fault diagnosis of the embedded system.

1. Introduction

With the rapid development of network information, the structure of computer networks is becoming more and more complex, the application fields are becoming more and more extensive, and people are increasingly dependent on the network. Therefore, the security and reliability of the network have become more and more important. As the main way to obtain the fault of the modern computer network and diagnose the system fault, the log uses the log information generated during the network operation to diagnose the fault, which has become a research hotspot of network fault diagnosis of the embedded system. Because there is a nonlinear mapping relationship between the fault type and the fault feature, the fault information has uncertainty. As the network architecture scale increases, a large amount of log information is generated during the running process,

and the fault feature extraction is very useful, thus increasing the difficulty of fault diagnosis [1]. In the existing fault diagnosis methods, the artificial neural network-based method has good fault tolerance, and it is not necessary for the domain experts to summarize the empirical rules of the diagnosed system from the domain knowledge or case, which is beneficial to overcome the knowledge based on the symbolic reasoning method. Obtaining bottlenecks, unlike symbol-based reasoning methods, which are sensitive to errors in rules or models, supports parallel computing. As parallel technology and hardware developments have become more widely used, they have been used to implement fault diagnosis [2].

Data mining is to find hidden, effective, valuable, and understandable patterns from a large amount of unordered data, to find useful knowledge, to derive the trend and association of time, and to provide users with decision-making

level decision-making support ability. The association rule mining algorithm based on the topology constraint performs hierarchical coding of each device in which each alarm occurs according to the hierarchical relationship between the network elements obtained by the established topology model. The connection relationship between the network elements reflected by the topology structure, combined with the propagation path of the fault, obtains the constraint condition of the association rule mining process. In the process of mining association rules, whether two or more items may be connected as a set of items is limited by such constraints. The interaction and mutual influence between network devices is the root cause of alarm propagation. Therefore, the patterns found in the alarm sequences sent by mutually influential devices are more targeted and more meaningful. If the relationship between devices is not introduced, the alarm of the related device cannot be filtered out, and a large number of meaningless modes are easily generated. Therefore, it is very important to introduce the constraints of network topology, and combining network topology constraints with data mining techniques can reduce the impact of noise and data loss by using constraints to find valuable rules [3].

Based on the summary and analysis of previous research results, this paper studies the network fault diagnosis of the embedded system method based on topology constraint and data mining. Firstly, by constructing a topology search algorithm, a fault diagnosis topology model is established, which eliminates the filtering of association rules without the relationship of topologies. Based on the collection of network alarm data, the quality of training samples is studied and related concepts are defined. The fault feature extraction algorithm is given, the hybrid system model is proposed, and the data mining model based on association rules is analyzed. The model algorithm is applied to the simulation experiment of network fault diagnosis of the embedded system. The algorithm in this paper provides a reference for further research on subsequent network fault diagnosis of the embedded system.

2. Methods and Principles

2.1. Topology Constraint. The topology model is an abstract view of the network structure, which hides other aspects that are not related to the topology relationship. Through it, this relationship can intuitively understand the connection relationship of each network element in the network, discover the characteristics of the network connection, and study the network structure and network performance analysis and other aspects are very helpful.

The network topology model can filter out irrelevant alarm data through the topology relationship in the preprocessing stage. For a certain alarm to be analyzed in the original alarm data, the network element that uses the alarm finds all network elements or most network elements that have a topology relationship with the network element in the topology model and extracts the connected network elements. It is stored in a data structure and in the form of network elements to form a cluster of network elements.

Under the framework of mean shift tracking, the target model and candidate target model are, respectively, written as follows [4]:

$$\hat{q}_u = C \sum_{i=1}^{n} k\left(\left\|x_i^*\right\|^2\right) \delta\left[b\left(x_i^*\right) - u\right],$$

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta\left[b(x_i) - u\right], \tag{1}$$

where x_i^* is the pixel in the target area, x_i is the pixel in the candidate target area with center y, k is the spatially weighted kernel function, and $b(x_i)$ is the histogram interval corresponding to the pixel at position x_i.

The coefficient is used as a measure to measure the similarity between two eigenvectors of the target region and the candidate target region:

$$\hat{\rho}(y) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u}. \tag{2}$$

The probability of crossover and mutation adopts an adaptive mechanism. As shown in (3) and (4), the probability of cross mutation is adjusted according to the individual fitness value so that poor individuals are more prone to cross mutation and good individuals are easy to save [5]:

$$P_c = \begin{cases} P_{c2} - \dfrac{(P_{c2} - P_{c3})(F' - F_{\text{avg}})}{F_{\text{max}} - F_{avg}}, & F' \geq F_{\text{avg}}, \\[3mm] P_{c1} - \dfrac{(P_{c1} - P_{c2})(F' - F_{\text{min}})}{F_{\text{avg}} - F_{\text{min}}}, & F' < F_{\text{avg}}, \end{cases} \tag{3}$$

$$P_m = \begin{cases} P_{m2} - \dfrac{(P_{m2} - P_{m3})(F - F_{\text{avg}})}{F_{\text{max}} - F_{avg}}, & F' \geq F_{\text{avg}}, \\[3mm] P_{m1} - \dfrac{(P_{m1} - P_{m2})(F - F_{\text{min}})}{F_{\text{avg}} - F_{\text{min}}}, & F' < F_{\text{avg}}, \end{cases} \tag{4}$$

where F_{max}, F_{min}, and F_{avg} are the maximum, minimum, and average fitness values of the population, respectively, F' is the larger fitness value of the two individuals performing genetic manipulation, and there are $P_{c1} > P_{c2} > P_{c3}$ and $P_{m1} > P_{m2} > P_{m3}$.

There is a certain model transformation relationship between the matching points in most images to be matched, and the topology relationship between several feature points in the same image is also approximately satisfied in the other image. For this purpose, the topology constraint can be performed by calculating the angle between the three feature points to purify the initial matching point. It is known that the coordinates of the three pairs of matching point pairs are $(x_1, y_1)^T$, $(x_2, y_2)^T$, $(x_3, y_3)^T$ and $(x_1', y_1')^T$, $(x_2', y_2')^T$, $(x_3', y_3')^T$, respectively, and the two angles p and q can be obtained by using the vector angle cosine formula [6]:

$$p = \arccos\left[\frac{(x_1 - x_2)\cdot(x_1 - x_3) + (y_1 - y_2)\cdot(y_1 - y_3)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\cdot\sqrt{(x_1 - x_3)^2 + (y_1 - y_3)^2}}\right],$$

$$q = \arccos\left[\frac{(x_1' - x_2')\cdot(x_1' - x_3') + (y_1' - y_2)\cdot(y_1' - y_3')}{\sqrt{(x_1' - x_2')^2 + (y_1' - y_2')^2}\cdot\sqrt{(x_1' - x_3')^2 + (y_1' - y_3')^2}}\right]. \tag{5}$$

For the above algorithm, a set of relatively stable correct matching points need to be found as the initial point pair of the calculation. Therefore, the key step of the algorithm is to find a set of such points because only one group needs to be found in all the initial matching points. The point fits the requirements of calculation, so algorithm can search it by increasing the set conditions, and the specific implementation method will be shown in the following network topology model.

2.2. Data Mining. Data mining technology is a database technology that extracts hidden, previously unknown, and potentially valuable knowledge and rules from vast amounts of data. The interest association rule directly indicates the recursive relationship between interests, and by simplifying the data in the buffer represented by the data model, it directly indicates the link relationship between the pages, which cannot directly reflect the degree of association between interests [7].

It is supposed that there are n learning samples, each with a set of observations $(x_1, x_2, \ldots, x_m, y_i^*)$ of $m+1$ parameters $(x_1, x_2, \ldots, x_m, y_i^*)\cdot(i = 1, 2, \ldots, n)$, $n > m$, and in practice, n is generally much larger than m, so as to ensure the accuracy and representativeness of the prediction results. The n learning samples of m parameters are defined as n vectors, that is, the expression of the learning sample is [8]

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{im}, y_i^*,), \quad (i = 1, 2, \ldots, n). \tag{6}$$

Let x_0 be the general form of a vector in $(x_{i1}, x_{i2}, \ldots, x_{im})$. The principle of the algorithms is the same, that is, to create an expression $y = y(x_0)$ and minimize its value:

$$\min(y) = \sum_{i=1}^{n} [y(x_0) - y_i^*]^2. \tag{7}$$

In addition, algorithms are to establish an equation $y = y(x_0)$ to maximize the classification interval based on support vector points to get the optimal separation line. The accuracy of the results obtained is different because of the different methods used by these algorithms.

The data are preprocessed by the analytic hierarchy process to obtain the weights W of different nodes, and the order S of each node is obtained. The lower side correlation function is set as [9]

$$h_{mn}(z) = \begin{cases} 0, & z \notin [z_n(1), z_n(2)], \\ \dfrac{z_{mn} - z_n(1)}{z_n(2) - z_n(1)}, & z \in [z_n(1), z_n(2)], \\ \dfrac{z_n(4) - z_{mn}}{z_n(4) - z_n(3)}, & z \in [z_n(3), z_n(4)], \end{cases} \tag{8}$$

where $z_n(1)$ is the minimum value of the sample data, $z_n(2) = z_n(3)$ is the number average, and $z_n(4)$ is the maximum number of data.

The preprocessed sample data are $Z = [Z(1)\ Z(2),\ldots,Z(M)]\ T$, and the information matrix $H_{M\times N}$ of (9) can be obtained via (8):

$$\mathbf{H}_{M\times N} = \begin{bmatrix} h & h & \cdots & h \\ h & h & \cdots & h \\ \vdots & \vdots & \vdots & \vdots \\ h & h & \cdots & h \end{bmatrix}. \tag{9}$$

By $q_{mn} = h'_{mn}\text{-}t_n + \varepsilon$ ($m = 1, 2,\ldots,M$ and $n = 1, 2,\ldots,N$), $t_n = \min(h'_{mn})$, the negative number is shifted to a positive near zero number, the obtained positive matrix $\mathbf{Q}_{M\times N}^n$ is calculated as follows:

$$\mathbf{Q}_{M\times N}^n = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1N} \\ q_{21} & q_{22} & \cdots & q_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{M1} & q_{M2} & \cdots & q_{MN} \end{bmatrix}. \tag{10}$$

The association analysis method in data mining technology can be used for association discovery, sequence pattern discovery, and same time-series discovery. For the sake of simplicity, the transfer relationship between interest association rules cannot be considered when establishing interest association rules. For this simple interest association rule association model, using association discovery analysis method is more appropriate.

3. Topology Modeling Based on Fault Diagnosis

3.1. Topology Lookup Algorithm Implementation. At present, all network topology models have some shortcomings, especially the topology-level modeling research at the router level is still in its infancy, and the current research results and methods have yet to be enriched and developed. As seen

above, as a typical representative of complex networks, the network topology has many basic parameters, such as node degree and its distribution, aggregation coefficient, median, and core number, but more importantly, new metrics to be discovered are metrics and network performance, such as synchronization performance and routing performance. The internal mechanisms of interaction between the two require further research by a large number of numerical simulations and empirical studies to provide a more in-depth qualitative and quantitative analysis of topology characteristics and metrics.

The network topology model lacks completeness for one or a few topology characteristics. In topology modeling, the existing research results show that model based on the design optimization principle is more consistent with topology structure than the scale-free model based on random principle. The network topology tree model used for topology lookup algorithm is shown in Figure 1. The network performance is also closer to the real network topology and the traditional scale-free model cannot be used to characterize the generation mechanism of network topology characteristics, such as robustness and vulnerability [10].

In a complex or untreated network environment, due to network attacks, hardware errors, and environmental obstacles, data are prone to distortion or error during transmission or positioning [11]. This is different from simple errors that are easy to occur in ordinary networks, and the data will seriously affect the positioning result. The algorithm is characterized by maximizing the correlation between the signal space and the physical space after the data transformation and still retains the local spatial topology information. Therefore, the points adjacent to the coordinates are in the transformed signal space, and this paper focuses on the research of positioning methods in this network environment. This kind of erroneous data with serious deviations can also be called the wild value, and its value is greatly deviated from the collected normal data. In other words, this outlier feature makes the erroneous data far away from other normal data points in the geometric spatial distribution, thus having a lower distribution density, so the motivation for topology modeling is in the same dataset. The lower the point, the lower the impact on positioning to achieve better robustness. Figure 2 is the topology analysis algorithm model flow.

The center position of the central core tracker is the initially selected center of the target area to be tracked; the target area of the start frame is manually given, the target area of the nonstart frame is obtained by tracking the previous frame image, and the border core tracker is the center position and is obtained by corner detection [4]. The corner point is a point where the gray level changes sharply in the image, and there are usually a large number of corner points at the boundary between the target and the background. The kernel tracking algorithm often uses the gray histogram to construct the target model, but the gray histogram is susceptible to image noise or changes, while the gradient direction histogram has the characteristics of being insensitive to the changes, which can be more robust. Target description is discussed, especially when the structure and
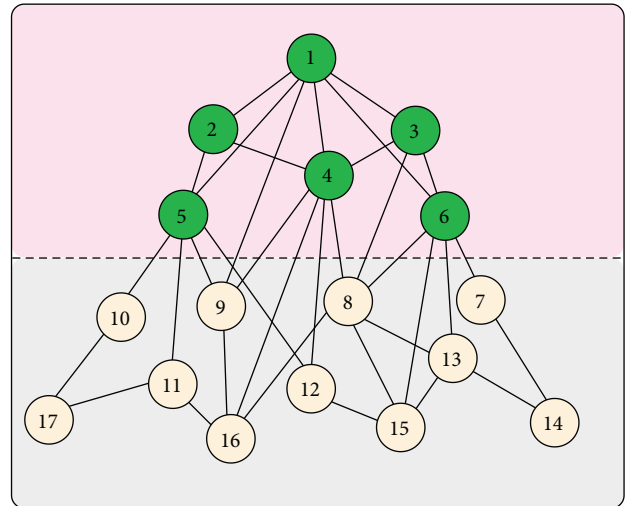
Figure 1: Network topology tree model used for topology lookup algorithm.

gradient information of the tracking target is rich. Using these nuclear trackers at the target and background, the constructed kernel histogram contains both the target information and the relatively stable background information, which can improve the validity of the target model.

3.2. Fault Diagnosis Topology Model. Fault diagnosis pattern recognition is an indispensable part of topology network fault diagnosis of the embedded system; the topology network can simulate the organization framework and construct a cognitive process to realize the fault handling. The working principle of this diagnostic method is to transmit the fault symptom to the system through the topology network and use the recognition model to classify the fault and obtain the diagnosis result. This type of diagnosis mainly uses the method of integrating the fault diagnosis data and, at the same time, training and practicing it. The data distributed in the topology network are expressed by means of computational translation, and finally, the fault diagnosis result is output. There are totally three algorithms which were compared in study, and they are diagnosis fault analysis system testing (DFAST), basic analysis system testing (BAST), and step fault analysis system testing (SFAST), respectively.

Topology networks can also integrate and collect ambiguous fuzzy fault information and integrate them into a complete fuzzy relational framework by means of mathematical logic such as functions. Different fault diagnosis topology model algorithm has different correct rates of retrieval under different window sizes, which has a significant impact on topology modeling (Figure 3). This framework can make unclear factors and avatars reasonable. In the scope, it provides a basis for network fault diagnosis of the embedded system. The diagnosis principle is mainly based on the integration of faults and the manifestation of form information, constructing a framework system belonging to the function, then focusing the fault factors and
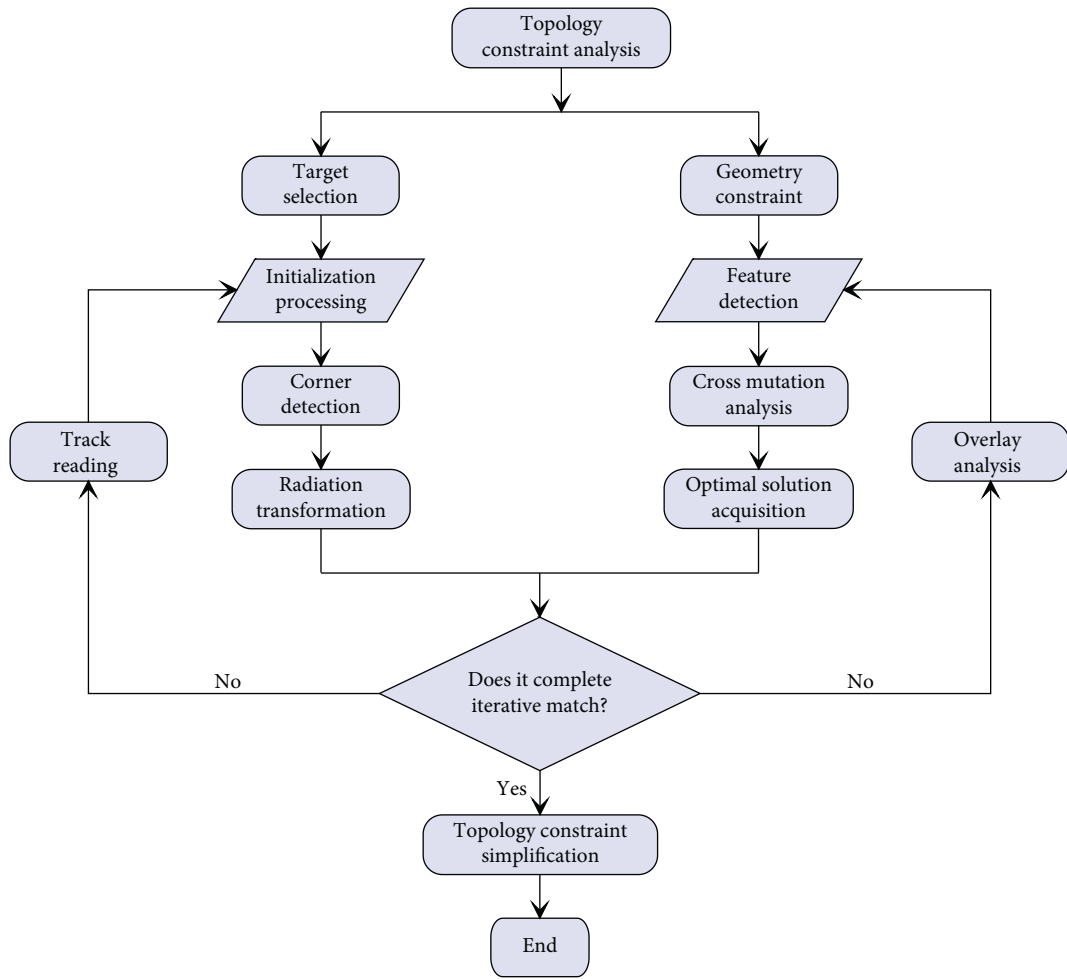
FIGURE 2: Topology analysis algorithm model flow.
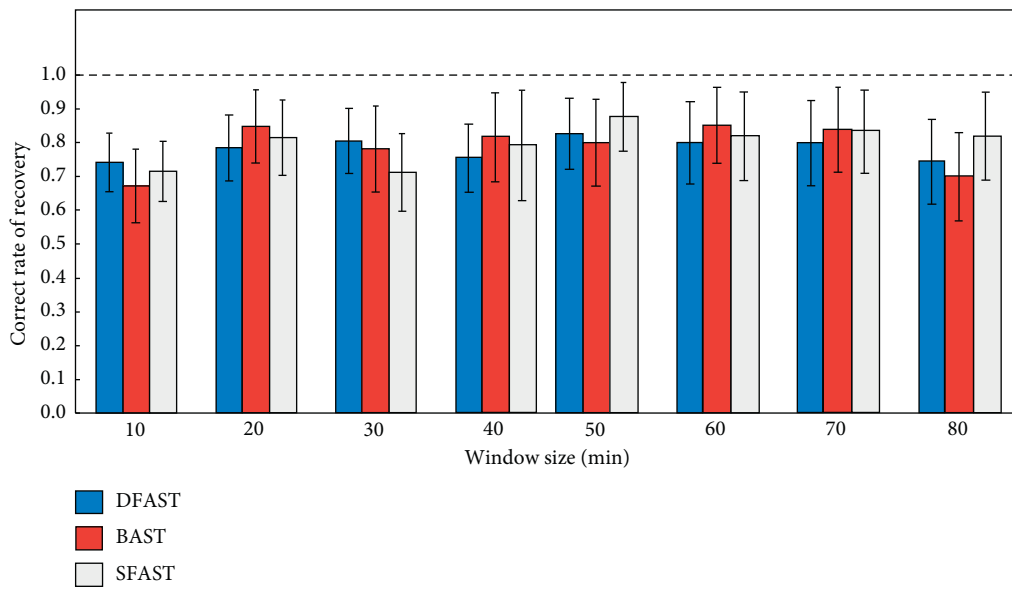
FIGURE 3: Correct rate of retrieval under different window sizes.

manifestations into a unified function framework, and using fuzzy relationship to carry out the fault factor category.

Network failures and signs of performance have strong uncertainties. The fuzzy relationship between the two makes it difficult to determine the cause of the two through common mathematical models, and the fuzzy logic network fault diagnosis of the embedded system method can be uncertain. The fault information is collected and integrated, and a relatively complete mathematical matrix model can be established through mathematical functions such as different functions. The establishment of the mathematical model can control the uncertain fault, and the symptom performance in a specific range and fault diagnosis provides a specific reference. The relationship between the cause and the symptom of the network system failure is extremely complicated and has strong randomness. Therefore, it is difficult to specifically locate the fault and analyze the cause by simply relying on the expert system. This kind of network diagnosis method can accurately judge the cause of the network failure and can bring some inspiration to the fault diagnosis of the maintenance personnel, but this method must be based on the establishment of a large fuzzy relational database and its intelligent learning ability.

The optimization of the network topology determines the weight of the network, thus transforming the problem of the neural network into a hierarchical decentralized optimization problem with the layer of neurons. The correction of the decision-making scheme is through neural network learning and rough set learning. The exchange between the two is improved until the rough set learning selects the decision rule with the least attribute composition to correctly divide all the training set samples. Although the training sample set is the knowledge given by the domain experts, including the knowledge formed by the fault condition attribute, there are inevitably incompatible situations, which will produce redundant samples. Training such samples will not only improve the correctness of decision-making but also reduce the learning efficiency of neural networks. On the contrary, each attribute of the fault symptom subset has different contributions to decision-making. The system needs to reflect the importance of different attributes in the learning process to improve learning efficiency [12].

4. Analysis of Alarm Data Based on Data Mining

4.1. Network Alarm Data Collection. Data extraction is the primary task of data analysis and the extracted data mainly includes alarm standardization data and network management system alarm data. It should be noted that the specific time zone should be specified before extracting data, and then, the time zone is determined by the network management system, and all alarm data in the data are extracted. After excluding the interference of the special data, the valuable data need to be filtered from the remaining data, and the interference alarm is removed. At the same time, the standardization field of the alarm information is set to the classification class and the weight class to judge the value of the alarm information. In order to avoid this problem, the predecessor proposed that, in the fault feedback phase, a

mobile sensor node is set in the detection area, and the mobile sensor node starts from the base station, collects the state information of all the sensor nodes by the mobile, and transmits to the base station. Finally, the selected valuable alarm information is integrated to form a network alarm data center.

If each node communicates with the base station, the node farther from the base station needs to communicate with the base station by means of the multihop node, which increases the energy consumption of the sensor node and is prone to error during transmission [13]. After obtaining all the data in the specified time period, the data should be cleaned, and the abnormal data, the missing data, the erroneous data, and the duplicate data, which affect the accuracy of the analysis, are eliminated, thereby ensuring the quality of the data analysis and realizing the accurate analysis of alarm information (Figure 4). Since the wireless sensor node has a certain communication range, it is not necessary to traverse each sensor node when performing fault feedback. The area where the sensor node is located can be divided and then the mobile sensor node can traverse each small area.

In the alarm association, the association analysis of the alarms can be used to obtain the strong association alarm rules with different support degrees and confidence levels. The minimum support degree and the minimum confidence level are used to help mine the root cause alarms to determine the alarms in the frequent alarm transaction set. In addition, we only specify the weight of the minimum support, the frequent item sets are filtered out, and all the confidences of the alarm associations under the frequent item sets are retained [14]. The purpose of this is mainly as follows: on the one hand, the frequent item sets in the massive alarm data can be filtered out as the target transaction set for determining the alarm weight by specifying the minimum support; on the other hand, all the alarm associations are retained. The confidence level can compare the relationship between all alarms' events in an alarm transaction set to determine the weight of each alarm event.

The core idea of the police weighting process of the big number theorem is to frequency the alarm weights, that is, according to the confidence level and alarm level of the alarm association rules, the alarm transaction set is taken as a set, and each alarm event is used as a random event through a large-scale simulation which calculates the frequency value of each alarm event occurring in this set as a weight. Alarm correlation analysis is a key issue in fault management, which helps network administrators delete a large number of redundant alarms, analyze the cause of the fault, and predict the occurrence of the fault. The traditional correlation analysis method mainly acquires relevant knowledge through experts, so it cannot meet the needs of network maintenance under the characteristics of large-scale, complex, and heterogeneous network, and the data mining method can make up for this deficiency.

4.2. Association Rule Mining Model. The association rules are different from the traditional association rules. The reliability of each item included in the requirements of the
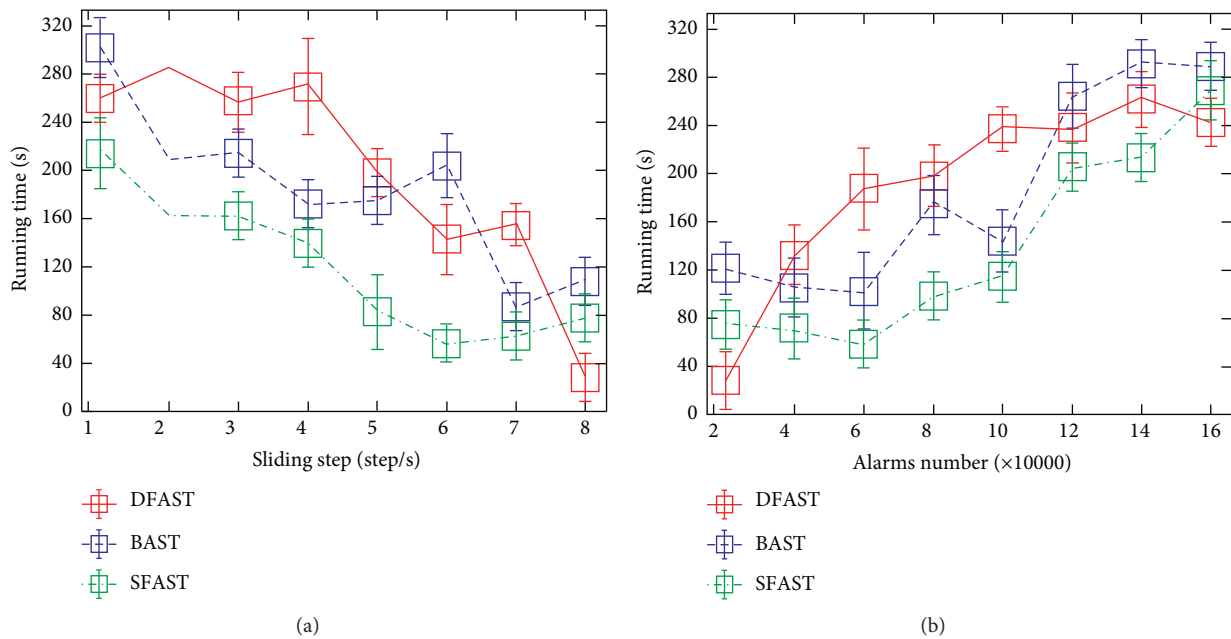
Figure 4: Relationship between algorithm running times and sliding step (a) and alarm number (b).

trusted association rules is in the same order of magnitude and does not care about the size of the entire rule. The redefinition of confidence can reflect the credibility of the rules, and traditional support can no longer be considered. For the mining of trusted association rules, the algorithm uses the adjacency matrix to generate two trusted sets and then uses the maximal group idea to generate all trusted association rules, thus avoiding scanning the database multiple times. These new metrics are designed to reduce the generation of false rules. Combining them into the generation process of frequent item sets can greatly compress the number of candidate sets generated and can mine strong intimacy association patterns. However, most of these methods are still based on algorithms, not only to scan the database multiple times but also to discriminate the interest of each candidate, so the time performance is low.

When constructing the sample set, a certain number of records before and after each fault occurrence time point are used as fault samples, and the nonfault time records are divided into several normal samples, and the normal sample records are the same as the fault samples. Firstly, the dataset is preprocessed, and the feature and constructor vector are extracted; then, the network model is built, the model parameters are configured, and the network training is performed. After the requirements are met, the network parameters are saved and the trained model is output. When there is new sample data to be diagnosed, the data are input into the trained network model, and the fault diagnosis result is fault or normal.

Depending on whether duplicates of the same dimension are allowed, multidimensional association rules can be subdivided into interdimensional association rules and hybrid dimension association rules, allowing dimensions to appear simultaneously on the left and right sides of the rules. Relationship between node and node degree, node failure

rate, and diagnostic accuracy are shown in Figure 5. Dimensional association rules and mining of multidimensional association rules also need to consider different types of fields, namely, category data and numerical data. For category data, the general association rule algorithm can handle it, while the logarithmic data need to be converted into category data for processing. Numeric fields are divided into predefined hierarchies; these intervals are all predefined by the user, and the resulting rule is also called the static quantity association rule [15, 16].

The algorithm is the most influential algorithm for mining a completely frequent item set and the algorithm has two key steps: one is to find all frequent item sets and the other is to generate strong association rules [17–19]. The algorithm is also a breadth-first algorithm, and the algorithm is based on the a priori algorithm. The first method of the algorithm uses the a priori algorithm when scanning the database. When scanning again, it is no longer scanning the entire database, but only scanning the candidate set generated last time. The scanning also calculates the support of frequent item sets to reduce the support, and the time to scan the database is to improve the efficiency of the algorithm; the fusion of algorithm produces algorithm. When the database is initially scanned, a priori algorithm is used. When the generated candidate set size can be stored in memory for processing, it will be transferred to algorithm until all frequent item sets are found.

The entire process only needs to scan the database twice, but the number of candidate sets generated is relatively large and the algorithm also adopts the idea of database partitioning. The database is divided into several partitions and marked at the beginning of each partition. In the process of scanning the database, a candidate set can be added to the marked points of each partition, and parallel calculation is performed when calculating the item set. Discovering
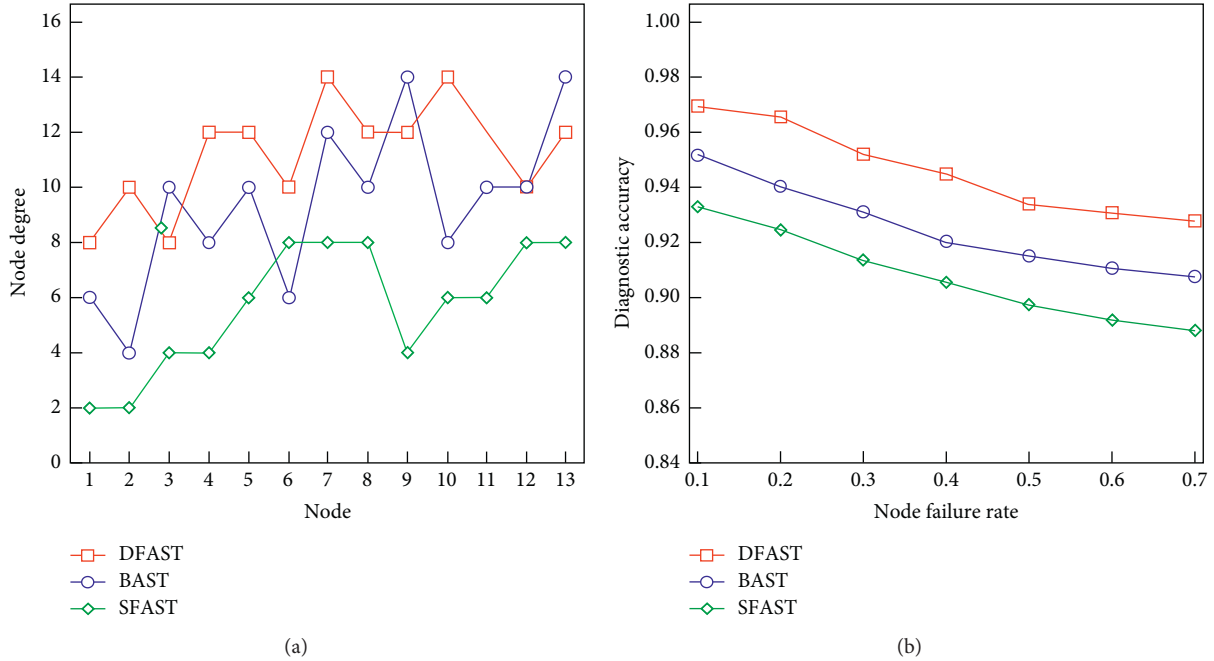
FIGURE 5: (a) Relationship between node and node degree. (b) Node failure rate and diagnostic accuracy.

frequent item sets is a key step in mining association rules. The algorithm also utilizes that the subset of frequent item sets is a frequent item set, and the super set of infrequent item sets is a nonfrequent item set, and this property effectively prunes frequent item sets.

5. Algorithm Simulation and Analysis

Since there is no link in the algorithm logic that combines the frequent patterns of item generation by the frequent patterns of the items in the classical algorithm, the tree structure is used for mining, that is, after the tree is generated, the frequent pattern can be generated in one step. Therefore, the network topology constraint is added after the algorithm generates all the frequent patterns, that is, after the algorithm digs out the frequent patterns, they are sequentially determined whether the network topology constraints are met, and then, the nonconforming patterns are deleted from the final frequent pattern set. Such an operation can avoid the infrequently frequent patterns being delivered to the user as the final output, which is more beneficial to the actual needs of the user.

The experimental results show that the network fault diagnosis of the embedded system method based on convolutional neural network can achieve better results than the traditional naive logistic regression and multilayer perception methods in the same training set and test set. Compared with the traditional network fault diagnosis of the embedded system method, the method based on convolutional neural network is simpler, no need to artificially extract the features in the text, and the network model configuration is simple, and the training data amount and model training time are reduced, and the matching calculation is reduced.

The topology search algorithm mainly constructs the classifier by discovering the association rules in the training set. Relationship between the remaining alarm number and the matching threshold value under different sliding steps is different from each other (Figure 6). In the construction process of the decision tree, the most time-consuming operation is to perform statistical calculation of the class distribution information on the dataset belonging to each nonterminal node and to split the dataset by using the splitting criterion. The discovery of association rules uses a classical algorithm that is effective for discovering association rules that are hidden among a large number of transaction records. However, when using it to discover classification rules, in order to prevent some rules from being missed, the minimum support is often set to zero. At this time, the algorithm cannot play its optimization role. As a result, the frequent sets generated sometimes cannot be in memory accommodated so that the program cannot continue to run. The method uses a group-based counting method to collect the category distribution information of various attribute value combinations in the training set and finds meaningful classification rules through the two thresholds of minimum confidence and minimum support.

In order to diagnose and analyze network faults based on data mining, it is necessary to first obtain information that reflects the relevant state of the network to a certain extent. Only the more completed and comprehensive the information of the relevant network state can be related. In actual network management, information exchanges between different management domains are relatively small due to security and other factors. The fault information is more included, which is more conducive to the discovery and
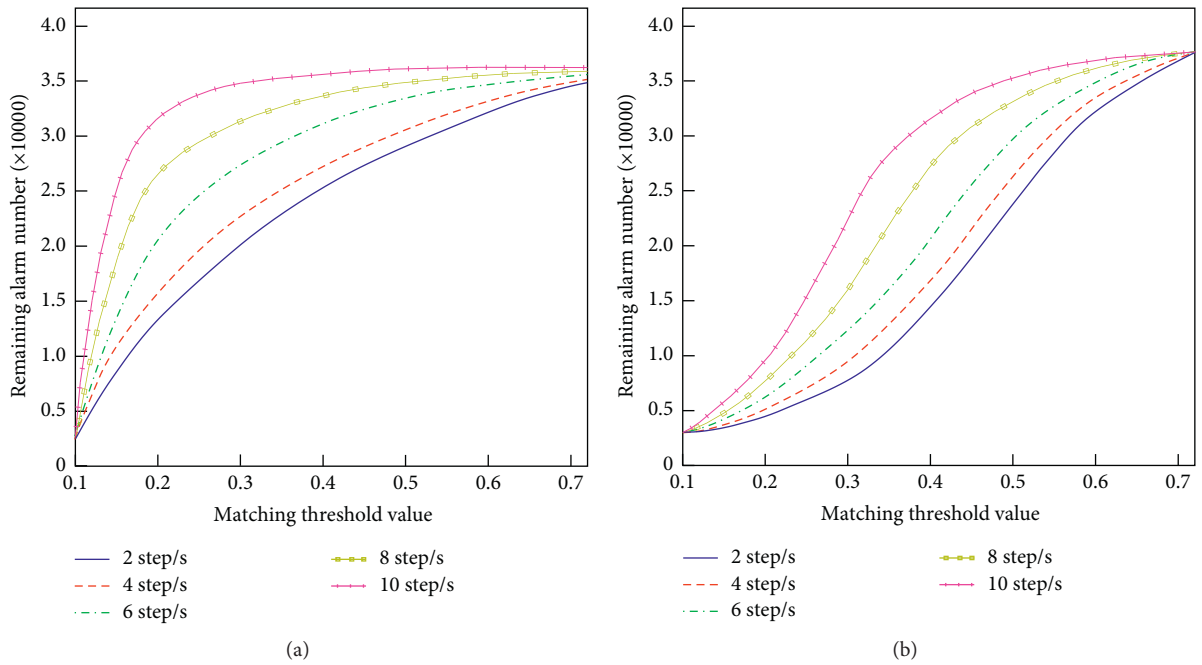
Figure 6: Relationship between the remaining alarm number and the matching threshold value under different sliding steps.

resolution of faults. If designers want to fully obtain relevant network information, they need to refer to the network management of the computer and related diagnostic tools for network failure and refer to how it provides relevant data. The statistical content in the network device can be obtained by performing certain accesses to related data items. After that, comprehensive analysis of the statistical contents of multiple related network devices can be carried out to realize the basic management of the network.

From the perspective of data mining, data warehouse is the platform for data mining implementation and data mining is an analysis and the decision-making method. Based on the principles of artificial intelligence, machine learning, and statistics, it analyzes and mines historical data based on data warehouses or data marts to find out the relationship patterns hidden in these data. It reflects the intrinsic nature of the data and makes a higher level of abstraction of the information contained in the data. The data mining process is divided into the following stages: data preparation, data selection, preprocessing, data reduction, target determination, algorithm determination, data mining, and pattern recognition and knowledge evaluation. The first four stages complete the data warehouse to prepare for knowledge discovery and five stages to mine useful knowledge. Data warehousing and data mining are inextricably linked and a typical data warehouse mainly includes data import, data warehouse and data mart, and access tools. The application of data warehouse and data mining technology in the field of remote monitoring and fault diagnosis greatly improves the efficiency and depth of data analysis and provides a

powerful means for intelligent fault diagnosis, especially fault prediction, so that remote equipment enjoys the expert level.

According to the mining technology, the associated rules reflect certain dependencies or some related knowledge between certain things. If there are two or more related relationships, any one of them can be certain predictions based on other attributes and use the four related parameters of confidence, support, minimum confidence, and minimum support to determine the next rule so that obstacle data that are not yet known according to this rule can be determined. Relationship between the remaining warning number and regular association number under different algorithms is shown in Figure 7. This standard also stipulates the items and types of data that must be used for the agent of the network, as well as the agents that allow the operation of the network. The extraction is performed and quantified accordingly to form a global network alarm model.

Centralized fault diagnosis is not a good solution, and data mining cannot be performed based on alarm information in real time. Researching distributed real-time fault diagnosis algorithms is of great theoretical value and practical significance for better reducing the time of system faults and improving the accuracy of fault location. Because different network equipment vendors do not agree on the alarm format and content and combine the characteristics of related network alarms, this paper first deals with the network alarm information, according to the characteristics of network alarms and related rules to carry out key attributes.
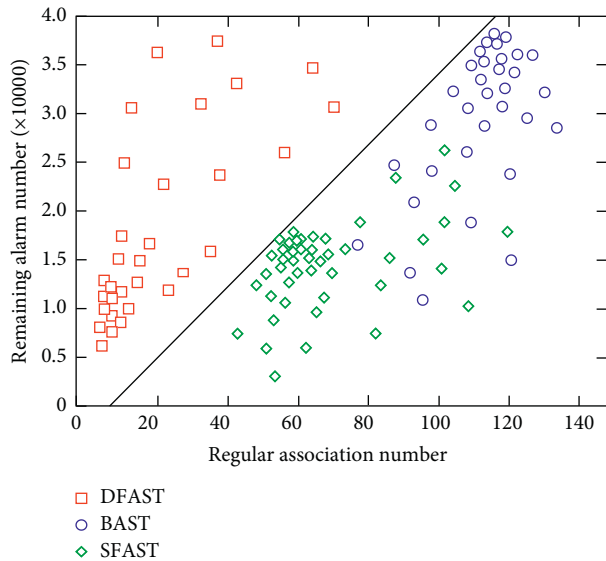
Figure 7: Relationship between the remaining warning number and regular association number under different algorithms.

6. Conclusions

This paper studied the network fault diagnosis of the embedded system method based on topology constraint and data mining. Firstly, a fault diagnosis topology model was established by constructing a topology search algorithm, which eliminated the filtering of association rules without topology relationship; the association rule-based data mining model was analyzed through the collection of network alarm data; the model algorithm was applied to the simulation experiment of network fault diagnosis of the embedded system and achieved good results. A new neural network algorithm based on rough set decision is proposed and the corresponding network fault is designed. A diagnostic model that manages all subnets and managed devices by providing a single network operation control environment is in a mixed network environment. The causes of noise data are data false alarms and measurement errors. When the fault occurs, the data of some variables of the system will fluctuate greatly, and some properties of the constructed network model will also change. The results show that correcting rage of retrieval varies from 0.65 to 090 under different window size; the running time of the proposed method drops from 310 s to 35 s during 1–8 step/s of the sliding step, while the node degree ranges from 8 to 14 and diagnostic accuracy ranges from 0.97 to 0.94; the remaining alarm number increases from 0.5 to 3.5 threshold value, while the regular association number is distributed in an interval of 40 to 140. The algorithm in this paper provides a reference for further research on network fault diagnosis of the embedded system. Moreover, further research studies should focus on the detailed influence of embedded system's inner structures on network fault diagnosis based on topology constraint and data mining.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] X. W. Li, "A network fault diagnosis model based on convolution neural network," *Software Guide*, vol. 16, no. 12, p. 4043, 2017.

[2] L. N. Hao, W. Wang, G. Y. Wu, and W. S. Wang, "Research on rough st-neural network fault diagnosis method," *Journal of Northeastern University*, vol. 24, no. 3, pp. 252–255, 2003.

[3] B. Xiao, Q.-F. Xu, Z.-Q. Lin, J. Guo, and C.-G. Li, "Credible association rule and its mining algorithm based on maximum clique," *Journal of Software*, vol. 19, no. 10, pp. 2597–2610, 2008.

[4] J. Su, B. H. Wang, and D. Z. Liu, "A multi-kernel tracking algorithm based on topology constraint," *Acta Electronica Sinica*, vol. 43, no. 2, pp. 353–357, 2015.

[5] Y. Y. Zhao, P. Yuan, Q. Ai, and T. G. Lv, "Improved genetic algorithm based on optimal configuration of PMUs considering topological constrains," *Power System Technology*, vol. 38, no. 8, pp. 2063–2070, 2014.

[6] X. P. Wang, X. Y. He, C. Q. Cheng, and L. Y. Fan, "An algorithm of image registration based on topology restraint of feature," *Geometrics and Spatial Information Technology*, vol. 39, no. 9, pp. 158–160, 2016.

[7] W. B. Xu and W. F. Zhang, "Applying data mining to web prefetching," *Chinese Journal of Computers*, vol. 24, no. 4, pp. 430–436, 2001.

[8] D. W. Li and G. R. Shi, "Optimization of common data mining algorithms for petroleum exploration and development," *Acta Petrolei Sinica*, vol. 39, no. 2, pp. 240–246, 2018.

[9] Z. Q. Geng, Y. T. Zhang, and Y. M. Han, "A fault diagnosis method for a Bayesian network based on AHP," *Journal of Beijing University of Chemical Technology (Natural Science Edition)*, vol. 44, no. 5, pp. 99–104, 2017.

[10] M. Zhou, J. H. Yang, H. B. Liu, and J. P. Wu, "Modeling the complex Internet topology," *Journal of Software*, vol. 20, no. 1, pp. 19–123, 2009.

[11] J.-J. Gu, S.-C. Chen, and Y. Zhuang, "Wireless sensor network-based topology structures for the Internet of things localization," *Chinese Journal of Computers*, vol. 33, no. 9, pp. 1548–1556, 2010.

[12] Q. M. Li, H. Zhang, and F. Y. Liu, "A new network fault diagnosis method based on RS-neural network," *Journal of Computer Research and Development*, vol. 41, no. 10, pp. 1696–1702, 2004.

[13] M. Y. Ma, Y. L. Zeng, T. T. Wei, and Z. D. Chen, "Fault diagnosis algorithm of WSN based on precondition of neighbor nodes," *Journal of Computer Applications*, vol. 38, no. 8, pp. 2348–2352, 2018.

[14] J. Yang and M. C. Lan, "Automatic switched optical network fault diagnosis algorithm based on alrm weighting," *Natural Science Journal of Xiangtan University*, vol. 40, no. 4, pp. 35–38, 2018.

[15] J. X. Bi and Q. S. Zhang, "Survey of the algorithms on association rule mining," *Engineering and Science*, vol. 7, no. 4, pp. 88–94, 2005.

[16] M. Chen, S. Lu, and Q. Liu, "Uniform regularity for a Keller-Segel-Navier-Stokes system," *Applied Mathematics Letters*, vol. 107, Article ID 106476, 2020.

[17] A. P. Wang, Z. F. Wang, S. G. Tao, and F. F. Yan, "Common algorithms of associaton rules mining in data mining," *Computer Technology and Development*, vol. 20, no. 4, pp. 105–108, 2010.

[18] H. Y. Liu, J. Chen, and G. Q. Chen, "Review of classification algorithms for data mining," *Journal of Tsingha University (Science and Technology Edition)*, vol. 42, no. 6, pp. 727–730, 2002.

[19] Y. L. Han, Y. Chen, S. Q. Cao, and M. L. Fang, "Remote service and fault diagnosis based on data mining," *Computer Integrated Manufacturing Sytems*, vol. 9, no. 2, pp. 158–162, 2003.

Retraction

Retracted: Teaching Mode Based on Educational Big Data Mining and Digital Twins

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Zhou and X. Wu, "Teaching Mode Based on Educational Big Data Mining and Digital Twins," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9071944, 13 pages, 2022.

Research Article

Teaching Mode Based on Educational Big Data Mining and Digital Twins

Xueyun Zhou [iD][1] **and Xinling Wu**[2]

[1]*School of Engineering, Guangzhou College of Technology and Business, Guangzhou 510850, Guangdong, China*
[2]*School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 510665, Guangdong, China*

Correspondence should be addressed to Xueyun Zhou; zhouxy@gzgs.edu.cn

Data mining technology has gradually become an important data analysis and knowledge discovery technology widely used in many modern industries. Data mining is a technique to find its regularity from a large amount of data by analyzing each data. It mainly includes three steps: data preparation, regularity search, and regularity representation. Data preparation is to select the required data from relevant data sources and integrate them into a data set for data mining; regular search is to find out the regularity contained in the data set by a certain method; regular expression is to be as user-readable as possible. The way of understanding (such as visualization) will represent the found patterns. This research mainly discusses the improvement of teaching mode based on digital twin-based education big data mining. Through the research on the basic principles of data mining and digital twin technology, the student evaluation tool module based on digital twin and the relevant data analysis tool module of students based on digital twin education big data mining are developed. Data mining is carried out from the data of student performance, personal basic information, and evaluation information to find the correlation between various factors, find the hidden laws, and provide support for teaching decision-making. This paper also solves the problem of frequent communication with remote databases according to the characteristics of the database data required by students and improves the efficiency and scalability of education big data mining technology based on digital twins. The goal of the virtual interactive system of the digital twin-based CNC platform is to have both three-dimensional real-time monitoring and remote control functions based on a three-dimensional virtual CNC panel. This research integrates the three-dimensional real-time monitoring and remote control of the virtual interactive system, analyzes the system operation process, develops the system interface, and improves the system sub-functions; it builds an experimental environment, conducts example tests on various functions of the digital twin platform virtual interactive system, and performs virtual interactions system performance indicators are analyzed. 60% of students believe that their innovation ability has been improved after the implementation of the digital twin teaching model; 50% of students believe that their self-evaluation ability has been improved. Applying digital twin's educational big data mining to student information management, university teaching evaluation, student performance analysis, and examination system, it has played a very good guiding role in improving the level of school teaching management.

1. Introduction

In the modern higher education management, the level of informatization has been improved year by year. With the widespread use of campus cards and the accumulation of data from major business systems over the years, a campus big data environment has been formed. It is mainly reflected in the characteristics of large-scale, multi-type, high-speed, and low-density value of student data. How to effectively mine student behavior data has become an important content to improve the level of student work informatization management. The data mining education model not only makes up for the shortcomings of face-to-face education and single online education in the past, but is also consistent with modern education reforms. In the information age, educators must break traditional education concepts, explore new education models and methods, promote education and education reform, improve education and education quality,

and promote students to develop in a better direction with the help of information technology.

By building a student behavior big data analysis system, it analyzes and processes the behavior characteristics of students in school, and applies it to the identification of students with financial difficulties in the school's student financial aid center. Through the reliability of massive data mining and the scientific nature of related algorithms, it changes the original mode of identifying work for students with financial difficulties from families, and implements the goal of improving the informatization of college students' work. The data mining education model mainly integrates online learning and offline learning through high-quality educational resources and intelligent learning platforms. In the process of mixed teaching, learners can not only communicate with classmates, but also communicate online at any time through the Internet. At the same time, students with fast learning progress can gain knowledge in advance through online learning, and students with slow learning progress can use network resources to further confirm and integrate knowledge to meet their learning needs.

The progress of virtual simulation resource construction urgently needs to be improved, the education system lacks top-level design, and the system construction standards are not unified. Schleich et al. believe that digitization of manufacturing industry promotes the application of complex virtual product model (called digital twin model) in all stages of product realization [1]. Li et al. believe that current airframe health monitoring often relies on deterministic physical models and ground inspections. He uses the concept of dynamic Bayesian networks. Dynamic Bayesian networks integrate physical models and various sources of contingent (random) and cognitive (lack of knowledge) uncertainty in crack propagation prediction [2]. Zhang et al. believe that new national advanced manufacturing strategies such as Industry 4.0, Industrial Internet and Made in China 2025 are introduced to realize intelligent manufacturing, and the number of new workshops in developed and developing countries keeps increasing. Under personalized design requirements, more realistic virtual models reflect the real world of manufacturing workshops, which is crucial to bridging the gap between design and operation [3]. Seay believes that for a critical infrastructure such as the national grid, the delay between identifying a cyber attack and deploying defenses can be disastrous. ORNL scientists are working to solve this problem with a platform called "digital twinning," which provides real-time simulations of the grid so that system distortions can be identified almost immediately. The grid is often called the largest machine in the world. It is a complex and interdependent network consisting of power generation, transmission lines for long-distance transmission of power, substations for lowering the voltage of public utility systems, distribution lines connecting communities, and equipment for converting power for customers [4]. Fei et al. believe that digital twin (DT) is one of the most promising enabling technologies for intelligent manufacturing and Industry 4.0. The importance of DTs is increasingly recognized by academia and industry. It has been nearly 15 years since the concept of DT was first proposed. To date, many DT applications have been successfully applied in different industries, including product design, manufacturing, forecasting, and health management. However, there has not been a review of DT applications in industry. In order to understand the development and application of DTs in industry, he comprehensively reviewed the DT research status of key components of DTs, the development status of DTs, and the main application of DTs in industry. He also outlined current challenges and some possible directions for future work [5]. Tao et al. consider digital twinning to be a new and rapidly developing technology for connecting the physical and virtual worlds, which has recently attracted wide attention worldwide. This paper presents a new product design method based on digital twin method. First, the development of product design is briefly introduced. Then the framework of digital Twin driven Product Design (DTPD) is proposed and analyzed. A case is provided to illustrate the application of the proposed DTPD method [6]. Lydon et al. believes that new and refurbished buildings will require significant improvements in energy performance according to Swiss and EU targets for 2050. The construction industry, however, has a poor record of improving productivity and efficiency. As other industries have demonstrated, digital approaches can reduce product development time and costs by using numerical simulations to reduce prototyping. In addition, digital twin is an extensive computational model of a product that is planned to be improved over its life cycle by leveraging operational data. He presents coupled simulation of thermal design of heating and cooling systems integrated with lightweight roof structures. The shape of the concrete roof structure is optimized to provide a low-energy building element that is thermally activated to provide space conditioning from a renewable source of heat. This work focuses on modeling methods used in the energy sector to support the development of digital twins of multifunctional building elements [7]. Virtual models can predict, estimate, and analyze dynamic changes by understanding the state of physical entities through perceptual data. The physical object responds to the optimization in the simulation. Digital twinning can optimize the whole teaching process through the network-physical closed loop.

Today, with the rapid development of information technology, the development of scientific theories has brought about the rapid progress of science and technology [8, 9]. The progress of science and technology has an impact on the field of education all the time, and changes the way of education transmission. The promotion and dissemination of resources has played a role in fueling the flames, benefiting more and more people. The online virtual simulation teaching experiment system designed and implemented in this article follows the principle of "combining virtual reality, complementing each other, and realizing reality without vitality." It actively explores the complementary advantages of various types of experiments in the experimental teaching design, combining virtual simulation, remote experiment, and optimized combination of physical experiments. It analyzes and evaluates the problems of low student performance, almost zero communication between

teachers and students, single learning style, and long feedback period in the digital twin virtual teaching model, and proposes corresponding solutions. It combines the analysis and evaluation results to design the corresponding virtual teaching interaction center, and completes the development, testing, and evaluation of the entire virtual interaction center, and provides a decision-making reference for relevant education departments.

2. Teaching Mode Improvement Research Methods

2.1. Virtual Simulation Teaching Platform. Digital twin platform can be divided into online learning feature analysis technology and virtual-real teaching space fusion analysis technology. Through the deep integration, mapping, and mirroring between "online teaching" and "online learning," it can inject new digital productivity into online immersive teaching [10]. First, deploy the developed experimental project to the business server. The uplink of the business server communicates with the user's client to transmit the user's operating parameters and the experimental data returned by the business server; the downlink communicates with the network management server, used to transmit the signaling of the service server and the experimental result data returned by the network management server. Second, the network management server regularly collects the status information of the registered LTE base stations through the heartbeat mechanism, and maintains the communication link with the base stations in the alive state; and through the connected computer that controls the mobile phone, it controls the switch of the mobile phone. Third, after the mobile phone is turned on, it connects to the LTE base station with the strongest signal around to establish a network channel. Remote users can select experimental items through the client computer and perform experimental control, including the configuration of base station parameters, adjustment of signal strength, initiating mobile Internet services, intercepting communication signaling, and other functions. The virtual simulation teaching platform is shown in Figure 1.

2.2. Three-Dimensional Digital Twin Podium Design. The combination of real-time modeling technology and virtual-real fusion technology, mainly through big data analysis, mining and modeling technology, carries out association rule analysis based on various algorithms on the association between virtual and real teaching spaces, so as to deeply analyze the relationship between virtual and real teaching spaces, and around the association, and then promote the integration of virtual and real teaching space.

The development of virtual platform interactive technology of 3D digital twin machine tool and multimedia blackboard control instructions is the basis of remote control technology of 3D virtual numerical control platform. By studying the interactive technology of the virtual CNC platform, it is possible to operate the virtual three-dimensional CNC platform to send commands. The development

of immersive teaching control instructions can directly operate the multimedia blackboard. Coupling the interactive commands of the 3D virtual control platform with the control commands of immersive teaching can realize the remote control of the multimedia blackboard. After the virtual platform interactive technology is implemented, the digital twin teaching class needs to be equipped with a virtual numerical control platform. The three-dimensional digital twin teaching class corresponds to the virtual numerical control platform one-to-one, and the multimedia blackboard is remotely controlled through the virtual numerical control platform. As the control equipment of 3D digital twin teaching, the virtual numerical control platform has two main functions. First of all, the virtual numerical control platform can directly send debugging instructions to the 3D digital twin teaching terminal, which can simulate the teaching explanation process and adjust the key content of teaching to verify whether the 3D twin teaching has qualified performance. Then, the virtual numerical control platform can send control instructions to the multimedia server, and the multimedia server parses the control instructions into corresponding operation instructions. According to the different immersive teaching, the multimedia server sends different operation instructions to realize the remote control of the multimedia blackboard.

2.3. Main Interface of the Virtual Interactive System. The digital twin podium is a real-time dynamic mapping of the real online teaching space. Therefore, the real-time acquisition, transmission, and dynamic update of data are of great significance to it. The main interface of the virtual interactive system is mainly divided into the real-time monitoring area of the center podium and the surrounding chart menu bar area [11]. At the top of the system, there is a more detailed sub-menu bar and real-time display of the current frame number of the system. The chart area on the left shows the exact position of the current platform, the number of people online in real time, the current teaching code program, and the layout of the lecture content in turn. Client users can either watch the lecture process of the podium in real time through virtual roaming, or check the accurate information of the podium through the menu bar. In addition to the main interface, you can view other information of the CNC platform and class through the menu bar. The menu bar provides the entire class layout diagram, and the client operator can view the entire class status through the class layout diagram, understand the distribution of the platform, and prepare for the three-dimensional virtual roaming. Or it can select the multimedia file that needs to be transferred from the client, and then send it by the virtual numerical control panel. Hiding the sub-functions of the virtual interactive system through the sub-menu can provide the client operator with a better sense of immersion.

Divide the projection interval T on the Y-axis into k equally, then the length of each sub-interval is [12]

$$l = \frac{F_{x\max} - F_{x\min}}{k}. \tag{1}$$
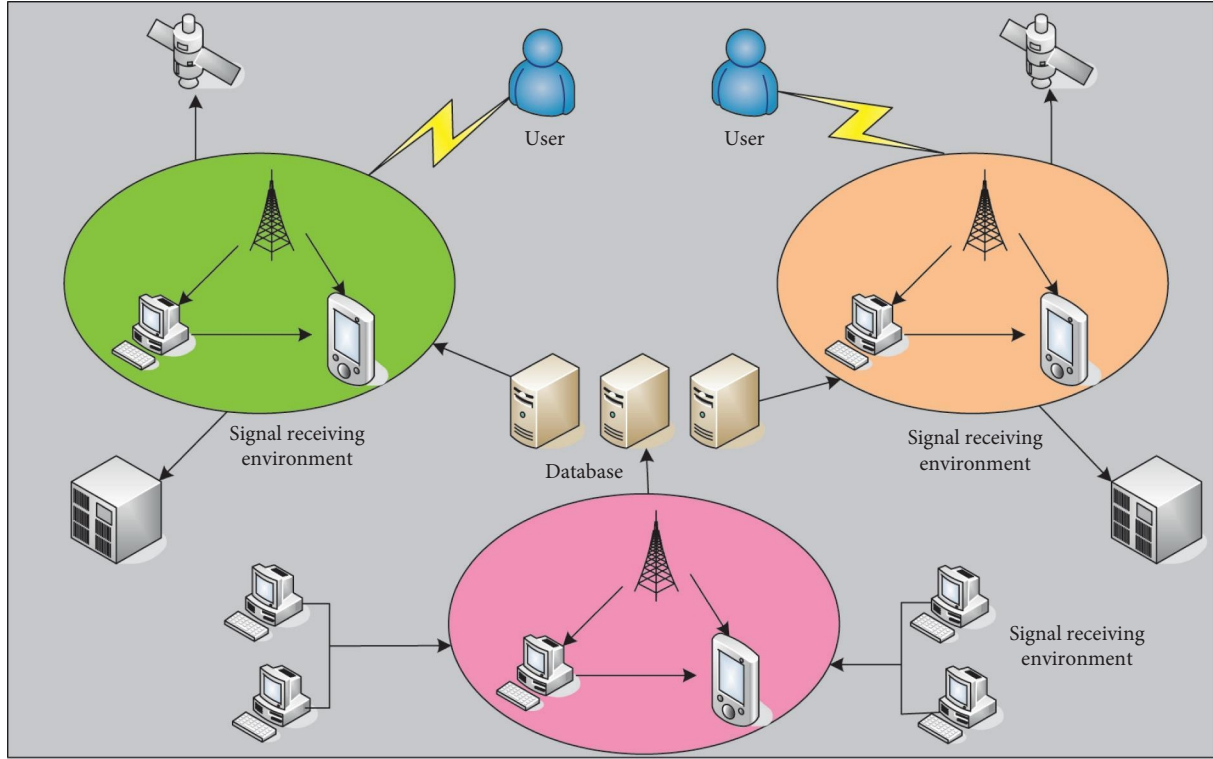
FIGURE 1: Virtual simulation teaching platform.

The points of the subset constitute the vertex subset S_D [13]:

$$S_D = (x, y, z)|F_{x\min} + \chi + f(x+1). \tag{2}$$

The mean value of the three-dimensional coordinate vertex distribution J [14]:

$$J = \frac{1}{3M} \sum_{i=1}^{m} (p + q + r). \tag{3}$$

The mean value of the peripheral vertex set can be obtained, and the covariance matrix C can be calculated [15]:

$$C = \frac{1}{3m} \sum_{i=1}^{m} (p_1 p_2 + q_1 q_2 + r). \tag{4}$$

C is a symmetric matrix, and its eigenvectors d_1, d_2, and d_3 are perpendicular to each other, and the three eigenvectors are unitized [16]:

$$\begin{aligned} d_1 &= \frac{d_1'}{|d_1|}, \\[6pt] d_2 &= \frac{d_2}{|d_2|}, \\[6pt] d_3 &= \frac{d_3}{|d_3|}. \end{aligned} \tag{5}$$

2.4. Network Communication Module Design. The network communication module is mainly realized through the Socket Channel Server and Socket Channel classes. The module is mainly divided into four parts: message server, message receiver, message parser, and message type registration. The network connection between the interactive teaching system and the teacher server is realized through the XmlClient class. XmlClient needs to configure the IP address and listening port number of the teacher server. If the connection is successful, then XmlClient will add the connected Socket to the connection list for subsequent data Receive and send. All interactive messages in the interactive teaching system are extended from the XmlMsg class, and applications can flexibly configure new message types by inheriting from this class. XmlMsgHandler is the corresponding message parsing class that matches with XmlMsg. All message types on the mobile terminal of the interactive teaching system need to be registered in XmlMsgFactory. In addition to actively connecting with the teacher server, the mobile terminal of the interactive teaching system can also establish an XML conversation service by using XmlServer. The XmlServer class can monitor the network and add the clients connected to it into the connection list.

2.5. Platform User Demand Discovery Function Design Based on Big Data Technology

(1) Platform users conduct learning activities and obtain portraits of current learning users. Platform users conduct learning activities through interaction with

related applications. When platform users have a demand for teaching resources, they will conduct data queries. The traditional information retrieval process is to match the user's query set Q with teaching resources U to obtain teaching resources. In order to realize the perception of user needs, first collect the information of each dimension of the user portrait at the current time t through the application and its events such as user records, workflows, and learning tasks related to learning activities. The ontology semantic description forms the current learning user portrait.

(2) User portrait matching is done, perceiving the demand for teaching resources. Calculate the similarity between the learning user portrait at the current time t and the historical application user portrait in the teaching resources of the integrated application user portrait, and through the process of demand discovery, demand screening and demand sorting, combined with platform user query conditions, the user's demand results are obtained.

(3) Optimize teaching resources and add teaching resources integrating user portraits. The results of the teaching resources required by the platform users are fed back to the users. The users select and evaluate the teaching resource sets pushed by the system, or re-search and query the teaching resources, and use the learning activities to obtain the teaching resource application results, and compare them with the current teaching resources. They learn the integration of user portraits and form teaching resources corresponding to new user portraits. The realization process of user requirement discovery is shown in Figure 2.

Identify and obtain direct user profile information. When platform users perform learning activities in order to complete corresponding learning tasks, the platform user's personnel identification, teaching background, and learning task identification indicate that it belongs to the direct acquisition of information.

Association defines user portrait information. Obtain the historical records of platform users through the association with the teaching resource system, including participating tasks, learning achievements, teaching resources concerned, retrieval data, and resource evaluation. Among them, manual semi-supervision and process management are required, which can infer the class where the platform user is located, the portrait of other users in the class, etc., so as to obtain a more comprehensive user portrait.

In the guest user portrait model tree UC2, the corresponding node is U', and the similarity calculation formula of the two nodes is [17]

$$\sim (U, U') = \sum_{m=1}^{n} w * \sim Q(U, U'),\qquad(6)$$

where $\sim(U, U')$ represents the node U and the corresponding node U' [18].

The formula for calculating the similarity of single-valued character attributes is [19]

$$\sim P(c_1, c_2) = \begin{cases} 1, & c_1 \cdot \text{value} = c_2 \cdot \text{value}, \\ 0, & \text{other.} \end{cases}\qquad(7)$$

For two different single-valued numerical attributes similarity:

$$\sim P(n_1, n_2) = \frac{\|n_1 - n_2\|}{|n_1|}.\qquad(8)$$

The similarity calculation of single-valued object attributes is equivalent to the nested calculation between user portrait elements [20]:

$$\sim(o_1, o_2) = \sum_{i=1}^{n} w_j \cdot \sim p(o_1, o_2).\qquad(9)$$

The similarity calculation method in the research process is [21]

$$\sim(s, p) = \max(\sim(s, s_i)|s \in p).\qquad(10)$$

Use $|PS|$ to denote the number of elements in the set, and the calculation formula for the similarity between sets is [22]

$$\sim(p_1, p_2) = \frac{\sum \sim(s, PS_1) + \sum \sim(s, PS_2)}{PS_1 + PS_2}.\qquad(11)$$

2.6. Specific Content of the Learning Evaluation Model for Digital Twin Teaching. Before the start of the course, first evaluate the students' learning goals. The purpose of evaluation is to determine whether the students' goals are reasonable, help students sort out their learning goals, and provide a good foundation and start for students' follow-up learning. At the input evaluation level, there are two aspects: the teacher's input and the student's input. The teacher's input includes the design and implementation of the mixed teaching model, and the establishment of a digital twin online learning platform.

The input of students mainly includes the emotional and attitude input of students in the learning process. This research mainly evaluates students. Therefore, only the input of students is selected as the main research content of input evaluation.

In process evaluation, the comprehensive ability of students is regarded as the main content of the evaluation. Teachers use the online learning platform to implement blended teaching, observe and analyze the learning behavior of students in the process of learning, and record and analyze the performance of students in activities and events.

In terms of achievement evaluation, the study effect of students is taken as the main research content, mainly to evaluate students' mastery of basic professional knowledge. Teachers learn the basic knowledge of students through various forms such as daily homework, group homework, individual independent homework, and final exams. The effect is comprehensively evaluated.
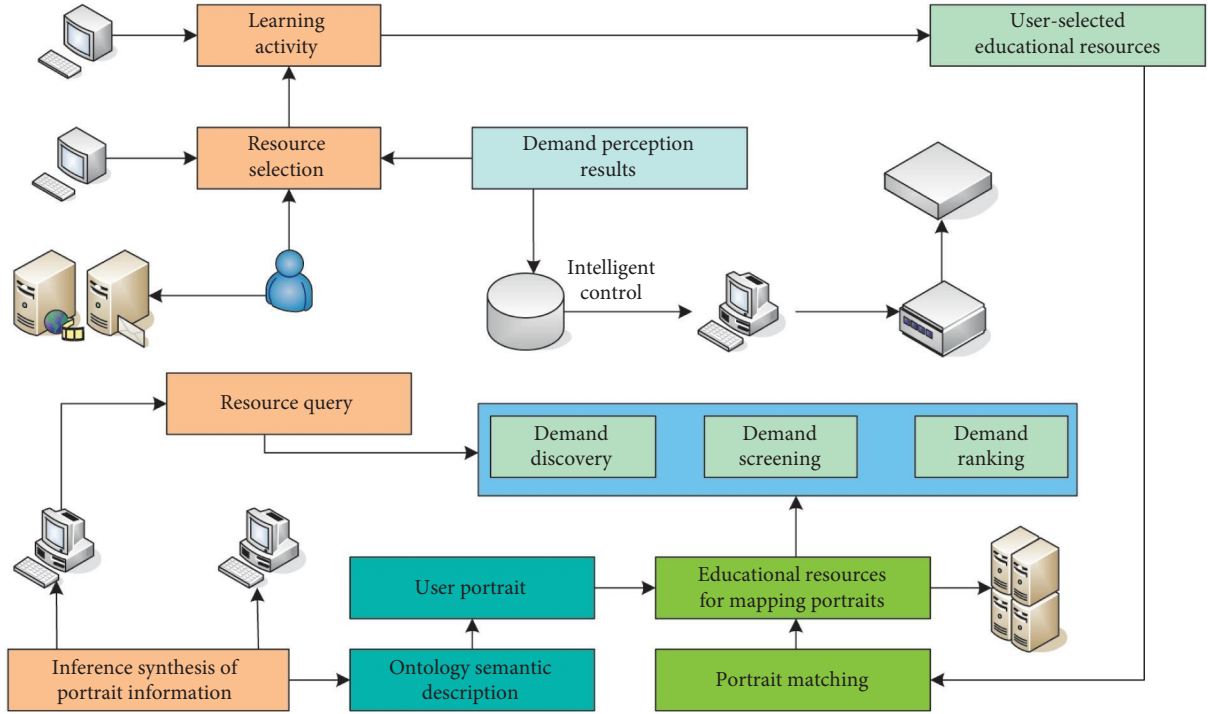
FIGURE 2: The realization process of user requirement discovery.

As a group, upload the work to the digital twin teaching platform. Groups can browse each other's works and evaluate other groups, and the scale for group evaluation will be uploaded to the platform by the teacher in advance. The component evaluation scale is shown in Table 1.

Teachers evaluate the students' personal assignments uploaded to the digital twin teaching platform, and score the students' daily assignments based on the personal assignment evaluation scale. The scale is also an important basis for evaluating students' learning effect and comprehensive ability. The personal work evaluation scale is shown in Table 2.

A general summary of the teaching resources that platform users need to query [23]:

$$C(QC) = \underset{i=1}{\overset{m}{U}} (mk(ef)). \tag{12}$$

The result set of educational resources formed after query and screening is as follows [24]:

$$MK(F) = \underset{i=1}{\overset{m}{U}} (mk(ef)), \quad m > 0, \tag{13}$$
$$MK = (MK_1, MK_2, \ldots, MK_n).$$

The big data demand results obtained are [25]

$$W(MK) = (\underset{}{\coprod} W_1(MK_1), \underset{}{\coprod} W_2(MK_2), \ldots, \underset{}{\coprod} W_N(MK_N)). \tag{14}$$

The set of teaching resource requirements is [26]

$$MK_{QC} = QC_P \cup QC_T \cup QC_O, \tag{15}$$
$$MK = MK_{QC} - MK_{ef}.$$

Digital twin learning satisfaction P_Q can be expressed as [27]

TABLE 1: Component evaluation scale.

First-level index	Points	Secondary indicators	Points
Topic selection	10	Reasonable topic selection	10
Content	40	Completeness	25
		Innovation	15
Express display	25	Language expression	15
		Gallery	10

$$P_Q = P_R \cdot \text{sim}_R + P_R \cdot \text{sim}_R + P_P \cdot \text{sim}_P + P_O \cdot \text{sim}_O. \tag{16}$$

The vector index of different learning objects is [28]

$$\overrightarrow{QC} = \{T_1, T_2, T_3, T_R\}. \tag{17}$$

The query vector of the learning process can be expressed as [29]

$$\overrightarrow{QC_g} = \{T_1, T_2, T_2 - T_3, T_3 - T_R\}. \tag{18}$$

3. Results of Educational Big Data Mining on the Improvement of Teaching Mode

The average score of the three rounds of action research is on the rise. Due to the step-by-step improvement of the teaching design, the number of people who get full marks has increased sharply. The higher degree of completion of classroom tasks also proves the rationality of teaching design and the improvement of teaching effect. The effect of teaching design is shown in Figure 3.

The sample students' learning objective evaluation grade scores are shown in Table 3.

TABLE 2: Personal work evaluation scale.

| Evaluation index | Grade | | | |
	Weights	Excellent (80–100)	Pass (60–80)	Unqualified (0–60)
Accuracy	0.4	Correct	Basically correct	Confusion, error
Innovation	0.3	Ideas and insights	Lack of novelty	Nothing new
Organized	0.3	Clear logic	Basically ok	No logic
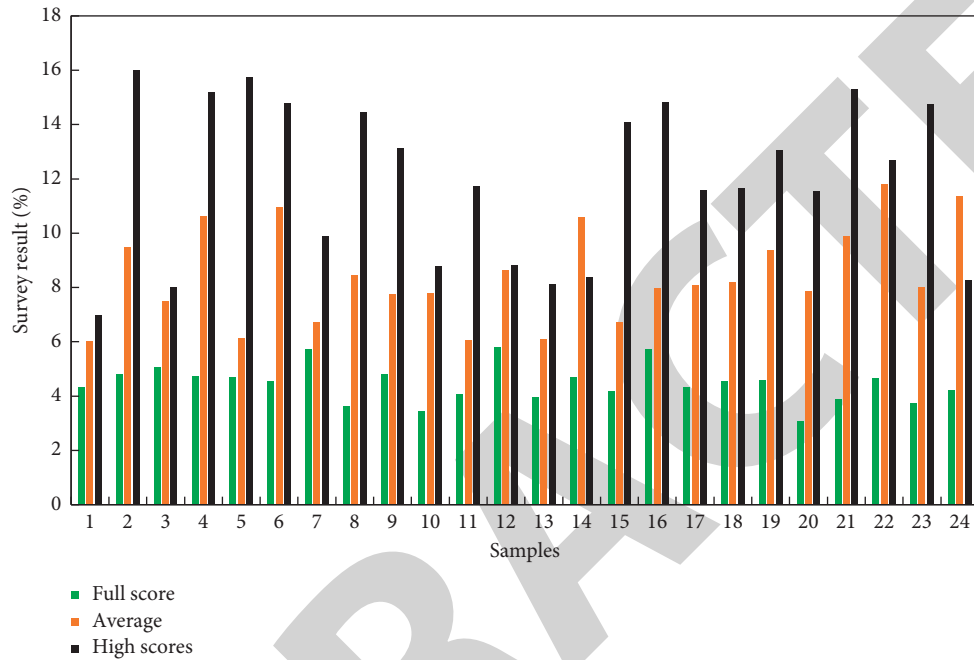
FIGURE 3: Teaching design effect.

TABLE 3: The learning objective evaluation grade scores of the sample students.

Learning target situation	Grade	Grade-corresponding score
The learning objectives are clear and feasible	A	95
The learning objectives are clear and feasible	A	95
The learning objectives are clear and feasible	A	95
The learning objectives are clear, and the feasibility of the learning objectives needs to be improved	B	85
The learning objectives are clear, and the feasibility of the learning objectives needs to be improved	C	75

According to the submission of students' works, the evaluation of students is shown in Table 4.

Calculate the average scores of comfort, adaptability, interactivity, sense of interest, and sense of immersion in the evaluation dimensions of students in the three classes of ABC. The comparison of each item is shown in Figure 4. It can be seen that with the update of the teaching design, the average scores of comfort, adaptability, interactivity, sense of interest, and immersion have shown an upward trend. Some have a large increasing trend, and some have a small increasing trend with a relatively high score.

In the course of teaching practice, the number of students who actively participated in the three rounds of action research was counted. The specific data are shown in Table 5. The first round of action research is the first time that digital twin teaching equipment has been brought to Class A classrooms. The number of students in Class A who are willing to try is relatively small. In the next two rounds of

action research, there are more people willing to participate in the digital twin teaching experience. After the first round of action research, taking into account the students' unfamiliarity with the new technology, the students will be specially trained in the simple operation skills of digital twin teaching before class. The students in the class behave more positively, and the classroom atmosphere is better.

Achievement rate is an intuitive standard that reflects the realization of teaching goals. In the process of teaching practice, the number of students who have completed the immersive exploration of digital twin teaching is deliberately counted. The specific data are shown in Table 6. Among them, only five students in Class A successfully completed the digital twin teaching immersion exploration, which indirectly shows that these students have a particularly strong ability to accept new things. The achievement rate of Class B and Class C is relatively high, which is partly related to the timely adjustment of teaching design. I feel the charm

TABLE 4: Evaluation of students.

Student ID	Accuracy (0.3)	Completeness (0.2)	Innovation (0.2)	Organized (0.3)
A	100	95	80	85
B	90	90	85	80
C	85	85	95	90
D	85	90	80	85
E	90	85	85	80
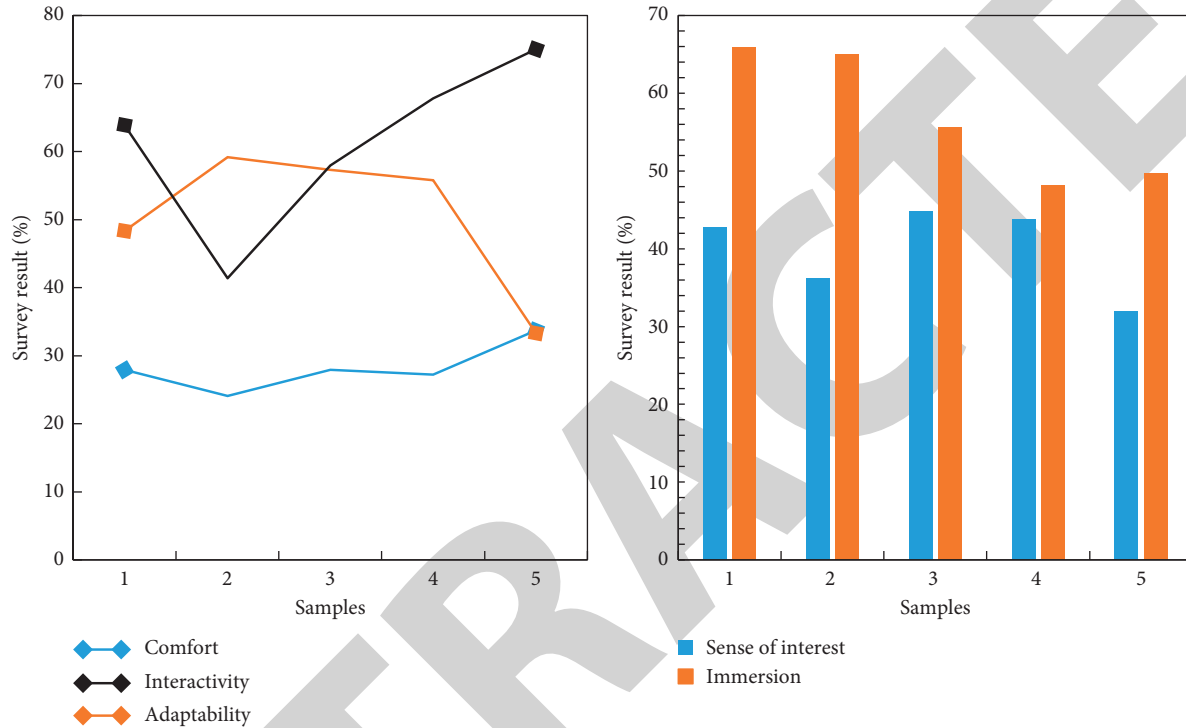
FIGURE 4: Each comparison.

TABLE 5: Number of students actively participating in the three rounds of action research.

Class	A	B	C
Total people	18	18	18
Number of active participants	8	15	17
Participation rate	44.44%	83.33%	94.44%

TABLE 6: Number of students who completed VR immersion exploration.

Class	A	B	C
Total people	18	18	18
Number of people reaching the target	5	12	17
Achievement rate	27.78%	66.67%	94.44%

of new technology more, and it is also an affirmation of the introduction of virtual reality technology into the classroom.

The performance test aims to test the performance of the mobile terminal of the interactive online teaching system when interacting with the teacher terminal. The timely response of interactive data has a great impact on the interactive experience, which needs to be paid attention to. The test is divided into non-optimized and optimized. The main test is the round-trip delay of interactive data between teacher and student equipment. Taking into account the frequency of interaction, the test mainly uses interactive whiteboard tools as the test target. Compared with the synchronization of courseware playback, the amount of synchronized data of the interactive whiteboard is larger, which can better explain the actual situation. First, consider the interactive data transmission round-trip time between the teacher and the student without closing the digital twin platform, as shown in Figure 5.

After closing the digital twin platform, the round-trip time of interactive data transmission between the teacher and the student is shown in Figure 6.

When the students answer the questions synchronously, the interactive teaching application has a small overhead, kept within 10%, which can be applied to more complex situations. The cost of interactive teaching applications is shown in Figure 7.

Among the students surveyed, more than 80% of students believe that their problem-solving skills have been improved after the implementation of the digital twin teaching model; more than 70% of students believe that their project practice skills, teamwork skills, and language expression skills have
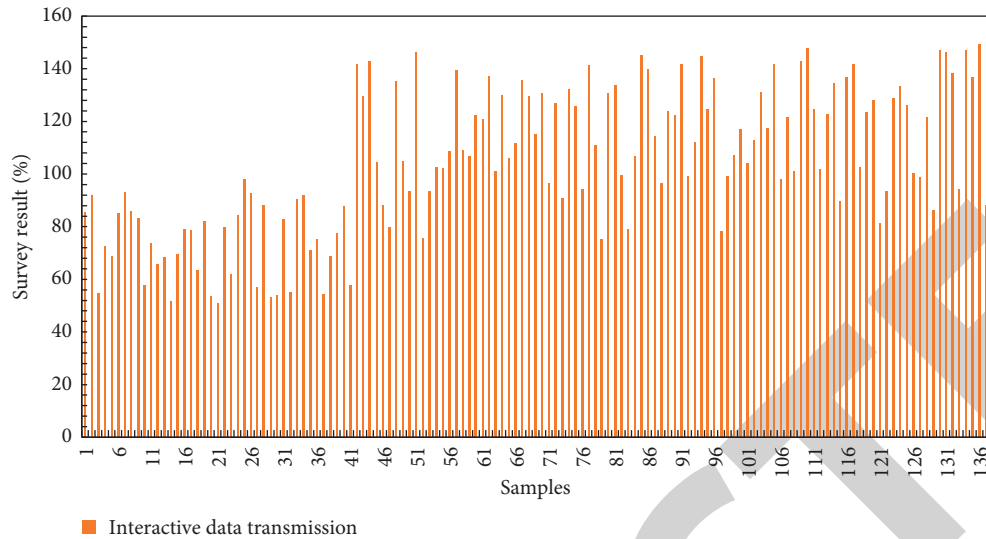
FIGURE 5: Interactive data transmission time between teacher and student without closing the digital twin platform.
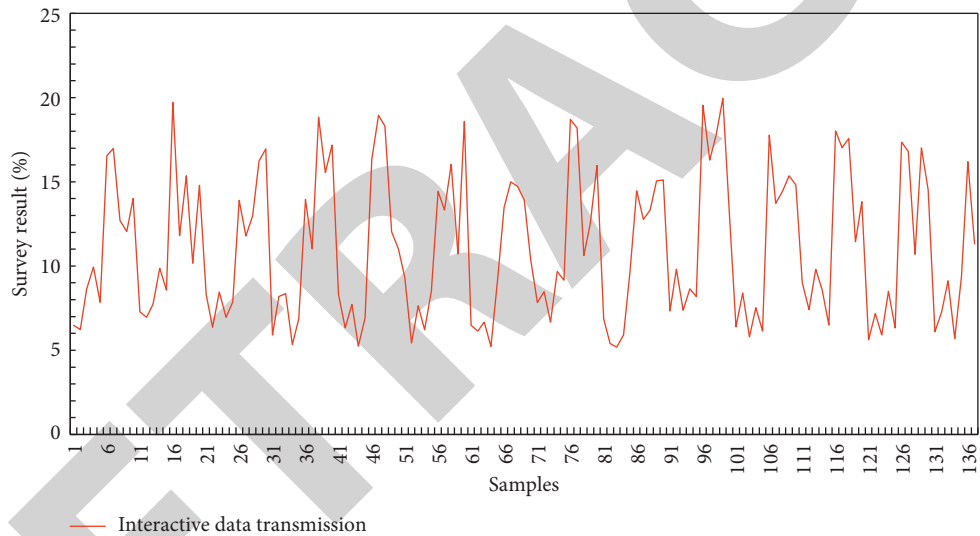
FIGURE 6: After closing the digital twin platform, the interactive data transmission round-trip time between the teacher and the student.

been improved; 60% of students believe that their innovation ability has been improved after the implementation of the digital twin teaching model; and 50% of students believe that their self-evaluation ability has been improved. The student ability evaluation is shown in Figure 8.

Students A and B have relatively high evaluation scores in terms of learning objectives, indicating that they have a clearer understanding of the course content before learning, but their learning attitude is at an intermediate level compared to other students and needs to be further improved. In terms of comprehensive ability and learning effect evaluation, further improvement is needed from the overall level. The overall level evaluation is shown in Figure 9.

4. Discussion

The teaching management department of colleges and universities needs to supervise and count the teaching quality of teachers, which requires the establishment of a scientific teaching evaluation system to carry out statistics and analysis of the status quo in the teaching quality process. Students are the objects of teachers' teaching courses, and they are also the key to improving teaching quality. Therefore, students' evaluation of teachers' classroom teaching has a very objective and realistic value. At present, many colleges and universities in China have adopted a teaching evaluation system based on student evaluation, teacher mutual evaluation, and self-evaluation. After many years of evaluation information feedback, many colleges and universities have accumulated a large amount of teaching evaluation data. The statistical results of the data serve as the basic basis for teachers' evaluation of awards, excellence, and job titles, and play the basic role of teaching evaluation data. However, with the development of information technology, simple data statistics can no longer meet the needs of university teaching management departments, because
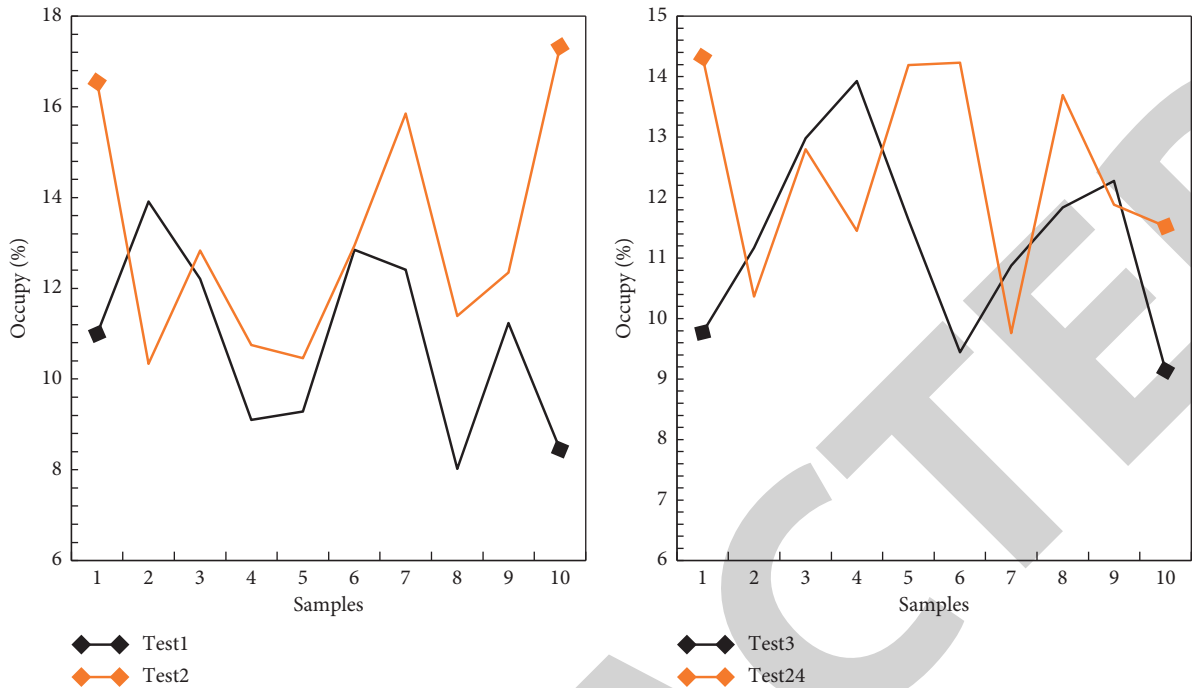
FIGURE 7: The cost of interactive teaching applications.
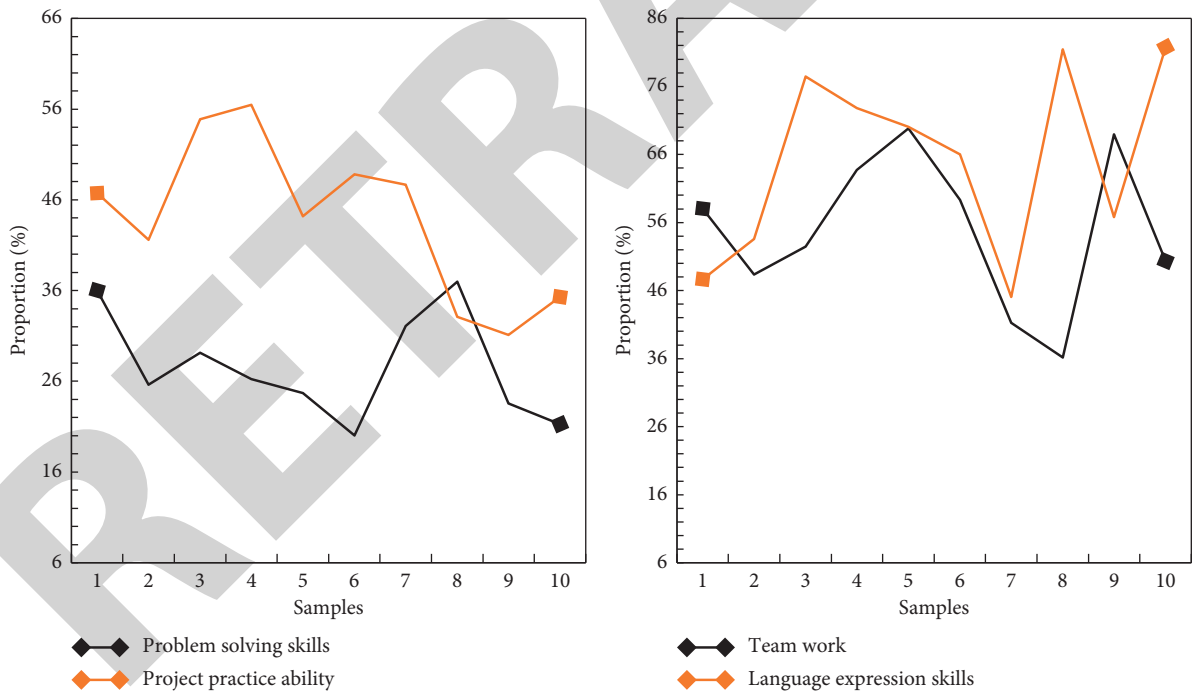
FIGURE 8: Student ability evaluation.

simple data statistics cannot find the potential teaching information that exists within the data, and more valuable data may be ignored by us. How to scientifically, rationally, and effectively use years of teaching evaluation data to guide teachers' teaching and improve teaching quality has become a very meaningful research topic.

Digital twin technology provides the architectural guidance of networked control and the direction of cyber-

physical system realization. With the help of digital twin technology based on informatization and networking, the development of the system can take advantage of the computing power and information fusion capabilities of the entire industrial network to give play to the advantages of new machine learning algorithms and data-driven algorithms. The leveling control and anomaly detection algorithms, as well as the realization of related algorithms, are
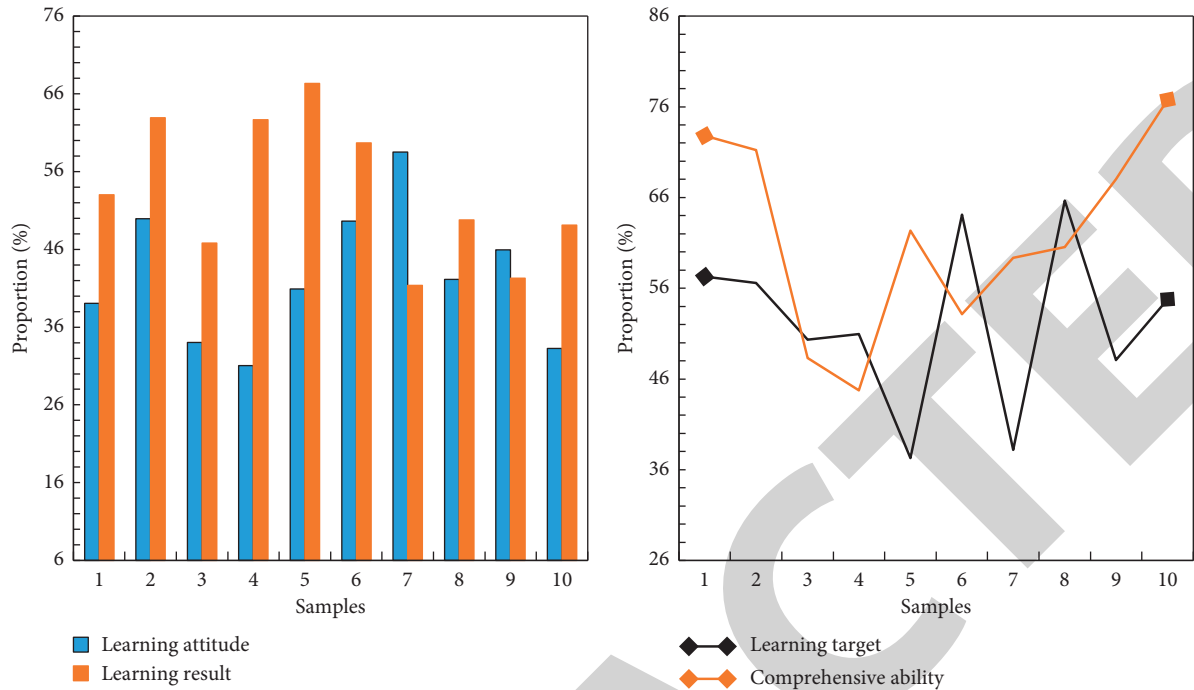
Figure 9: Overall level evaluation.

further studied and improved, combining information from different fields to achieve intelligent and networked control.

Subjective weight distribution requires the induction and extraction of human experience before the algorithm is executed. Although human experience has a strong individual dominance, the experience accumulated over the years cannot be replaced by any tool for induction and summary, and the value of human experience cannot be underestimated. The weight distribution based on objective conditions is actually the weight distribution based on objective data. It mainly analyzes the actual data that actually exists according to a certain algorithm, and finally obtains the weight value of the indicator based on the objective data. This method is based on objective facts. In order to analyze the data based on the facts, the results must be consistent with the objective facts. The objective weight determination method has a strong ability to persuade the data, and the results are reasonable. In the era of big data today, in the era of rapid economic and technological development in China, our focus has slowly shifted from food, clothing, housing, and transportation to education issues. We are in line with the world, and we want to accept better and more cutting-edge technologies. But what everyone can do in today's society is to perform their duties with full responsibility and help improve the way we are educated and the environment. This requires each of our students to face each answer sheet of teaching evaluation with a positive and serious attitude, and to attach their own opinions on classroom teaching after completing the selection of teaching evaluation indicators. At the same time, it is also required that educators in our schools and even the country can use the existing technical conditions to systematically analyze the real and effective teaching evaluation data, and combine the results of teaching experience and mathematical statistics to make an

adjustment, an improvement, so that all enjoy the power that technology brings to education.

Three-dimensional real-time monitoring based on digital twins requires high enough real-time data. Remote monitoring in digital twins is the basis of course early warning and remote control. In the past, the delay of monitoring systems is often at the second level, which cannot meet the requirements of remote real-time monitoring in digital twins. Three-dimensional real-time monitoring based on digital twins requires the graphics processing speed to be fast enough. The essence of the digital twin model of the virtual interactive class is the superposition of multi-dimensional elements in computer graphics. The digital twin model of the virtual interactive class is highly complex, the monitoring scene modeling is huge, and the data update speed is very fast.

The database is the basis of the entire system, and its quality is directly related to the realization of the system. In traditional education management and education concepts, education statisticians must meet the requirements of data mining, use different spreadsheets to handle the traditional data management of the new student data classification system, and process the data. This is a new design requirement of the database, which requires mutual conversion and spreadsheet functions. Due to the realization of global information resource sharing, the surge of information, and the improvement of information demand, traditional information extraction methods cannot meet the practical needs, so data mining technology is proposed. Data mining technology is used to extract and analyze large amounts of data to find hidden clues and provide a more valuable basis for further decision-making. Data mining is based on the database for research. The ultimate goal of the researcher is to combine the database and data mining to

make the final decision. In today's high-tech development, the scope of application of data mining is becoming wider and wider. It is not only suitable for university management, but also promotes the further reform, improvement, and development of school management, and also provides a good way for managers to make correct decisions. The educational management methods of basic schools have also been improved [30].

Generating a data warehouse is mainly about interface design and data loading interface design to ensure that the data extraction, transformation, and integration from the operational environment to the data warehouse can be completed and accurately completed with flexibility and variability. When loading data, ensure the order of data loading, filter invalid and wrong data, and manage data granularity. Design the interface from the database system to the data warehouse system to complete data purification and integration other than data extraction. The design of a data warehouse is an iterative process. As the process progresses, user needs are constantly clarified and the data that users really need are determined.

The national plan is based on education, and education is not a trivial matter. Teaching is different from business. Because there is no huge profit margin and long input-output time period, the application of high-tech in the teaching field is relatively backward and low in effectiveness. Appropriate methods are adopted to attract more scholars to use high-tech to transform teaching. It is a new problem facing the education system. Perform correlation analysis on students' personal basic information and results, various evaluation indicators, and project information that users are interested in, and look for the impact of students' personal basic conditions on student performance and evaluation indicators, as well as the interrelationship between results and various evaluation indicators. Data mining techniques are used to find the final result [31].

5. Conclusion

Although traditional face-to-face classroom teaching can effectively solve the above problems, it cannot meet the requirements of students in individualized learning and autonomous learning. In addition, traditional face-to-face teaching is mainly based on teachers, and students are in a passive position, which is not conducive to students' independence and innovation. The introduction of educational data mining technology into the digital campus system not only improves the scientific nature of education management, but also enhances the effectiveness of the digital construction of education. This system provides teachers and students with a digital twin virtual online interactive teaching platform that is easy to interact, manage, and operate. At present, the entire system has a large amount of data storage, including teacher data and student data. The education management information system only has simple query, statistics, and print functions for such a huge amount of data. If we can dig and analyze these data in depth, and discover some kind of internal connection, we will get some beneficial suggestive information. This article first made the

overall design of the system, then designed the specific functions of the system, and then showed the experimental results of the system. The research in this field has certain potential, and the efficiency of digital twins in this article needs to be further improved.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] B. Schleich, N. Anwer, L. Mathieu, and S. Wartzack, "Shaping the digital twin for design and production engineering," *CIRP Annals*, vol. 66, no. 1, pp. 141–144, 2017.

[2] C. Li, S. Mahadevan, L. You, C. Sergio, and W. Liping, "Dynamic bayesian network for aircraft wing health monitoring digital twin," *AIAA Journal*, vol. 55, no. 3, pp. 1–12, 2017.

[3] H. Zhang, Q. Liu, X. Chen, D. Zhang, and J. Leng, "A digital twin-based approach for designing and multi-objective optimization of hollow glass production line," *IEEE Access*, vol. 5, pp. 26901–26911, 2017.

[4] S. Seay, "Seeing double: digital twin for a secure, resilient grid," *Oak Ridge National Laboratory Review*, vol. 52, no. 2, pp. 34-35, 2019.

[5] T. Fei, Z. He, A. Liu, and A. Y. C. Nee, "Digital twin in industry: state-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.

[6] F. Tao, F. Sui, A. Liu et al., "Digital twin-driven product design framework," *International Journal of Production Research*, vol. 57, no. 11-12, pp. 3935–3953, 2019.

[7] G. P. Lydon, S. Caranovic, I. Hischier, and A. Schlueter, "Coupled simulation of thermally active building systems to support a digital twin," *Energy and Buildings*, vol. 202, pp. 109298.1–109298.19, 2019.

[8] F. Zhu, C. Zhang, Z. Zheng, and A. Farouk, "Practical network coding technologies and softwarization in wireless networks," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5211–5218, 2021.

[9] S. Dahmani, M. Gabli, E. B. Mermri, and A. Serghini, "Optimization of green RNP problem for LTE networks using possibility theory," *Neural Computing & Applications*, vol. 32, no. 8, pp. 3825–3838, 2020.

[10] S. Sun, M. Kadoch, L. Gong, and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," *IEEE Network*, vol. 29, no. 3, pp. 54–59, 2015.

[11] H. Jang, S. Hao, P. M. Chu, P. K. Sharma, Y. Sung, and K. Cho, "Deep Q-network-based multi-criteria decision-making framework for virtual simulation environment," *Neural Computing & Applications*, vol. 33, pp. 10657–10671, 2021.

Research Article

Regional Economic Prediction Model Using Backpropagation Integrated with Bayesian Vector Neural Network in Big Data Analytics

Qisong Zhang [iD],[1] Lei Yan,[2] Runbo Hu,[3] Yingqiu Li,[1] and Li Hou[4]

[1]*School of Information and Business Management, Dalian Neusoft University of Information, Dalian 116000, Liaoning, China*
[2]*Business School, Zhejiang Wanli University, Ningbo 315000, Zhejiang, China*
[3]*China Business Executives Academy, Dalian 116000, Liaoning, China*
[4]*School of Economics and Management, Ningbo University of Technology, Ningbo 315000, Zhejiang, China*

Correspondence should be addressed to Qisong Zhang; zhangqisong@neusoft.edu.cn

Forecasting economic growth is critical for formulating national economic development policies. Neural Networks are a type of artificial intelligence that may be used to model complex target functions. ANN (Artificial Neural Networks) are one of the most effective learning approaches now available for specific sorts of tasks, such as learning to understand complex real-world sensor data. This paper proposes the regional economic prediction model based on neural networks techniques. Bayesian vector neural network (BVNN) is integrated with backpropagation (BP) model. The database has been collected based on the economics of particular region which has been extracted and classified using knowledge-based computer analysis by neural networks. Discretization, reduction, importance ranking, and prediction rule are attributes considered here. Then, as the input training sample, feed extracted important components into the NN. This strategy enhanced the training speed and prediction accuracy by reducing structure of NN. WEO, APDREO, and AFRREO are the dataset and FWA-SVR and LSTM are the existing method taken for comparison. For the WEO dataset, 97% of GDP and 98% of accuracy are produced. For APDREO dataset, 92% of accuracy and GDP of 97% are obtained. For AFRREO dataset, 98% of accuracy is produced. The neural network can tackle nonlinear problems, according to experimental data, and the technology has been proven to be successful and viable with high accuracy. For practical application, the model has a good reference value. The proposed model reduces error by increasing the convergence rate and accuracy for each dataset.

1. Introduction

In evolution of human society, economic gain is a critical topic [1]. Based on surroundings, economic system is constantly influenced, which is also a nonlinear system. Additional interference variables, as well as different reasons that lead to ambiguous and inadequate historical information required for macroeconomic modelling [2], act directly on entire development of macroeconomic scheme process. As a result, the traditional BP_NN regional economic trend prediction technique cannot effectively correct weight in analysis process of vast data and operation efficiency is low. A typical BP_NN learning process includes two stages:

forward data propagation and backward error propagation. Input signal is analysed by neurons in input and hidden layers before being sent to output layer to display result [3]. It uses reverse propagation procedure if intended output is not accessible at output layer. Along originally linked path, there is a difference between network output and returned actual value. Modifying weights of neurons at each layer can lessen inaccuracy. Return to forward propagation process and continue until error is less than specified value. As a result, an enhanced BP_NN based approach for predicting regional economic trend is created, providing a solid foundation for regional economic development [4]. Forecasting regional economic activity is an increasingly

significant component of regional economic research. Regional economic predictions can directly aid local, subnational, and national policy makers as well as concerned business executives. Both policymakers and business executives need accurate predictions of key economic aggregates, such as employment, income, and output for medium-long term planning purposes. Regional economic activity forecasts have been used to explain macroeconomic forces, including predicting the stock market and cyclicality of national labour market movements. Furthermore, international investors and multinational agencies engaged in megaprojects at a regional level also need accurate predictions for investment planning reasons [5].

There are numerous statistical tools and models available today for assessing and developing integral indicators of quality of life. Most methods, however, have limitations in terms of applicability and do not function with incomplete or "noisy" data. As a result, neural network technologies and intelligent information systems based on them are presented as an alternate strategy for handling semistructured and normalised challenges of analysing quality of life and developing forecasting models [6–8]. Savings, housing, socioeconomic features of the population, the dynamics, and differentiation of these indicators in the context of demographic groupings and regional economies are the subjects of several studies [9]. Systems of absolute and relative indices are established in these studies to characterize the degree of provision and fulfilment of the demands of various groups of people. Authors use ways to identify the concepts of "cost of living" and "quality of life" in several models of the standard of living.

The cost of living is defined in classical economics as the value of a set of commodities that is determined, first and foremost, by the level of food and service costs, as well as changes in types of income, consumption habits, tax legislation, and demographic indicators of population groupings. Consumer basket and commodity bundle, as well as living wage [10], are the major estimated rates. The presence of structural shifts and the interaction between different sets of indicators, as well as nonlinear dependencies between indicators, are characteristics of time series of socioeconomic indicators [11–13]. Therefore, linear models are not always able to give qualitative results. In this regard, a hybrid approach based on the construction of an ensemble of hybrid models is promising, in which, along with regression models, models based on neural networks, decision trees, and neurofuzzy networks are used. In recent years, in many works, the hybrid approach has been developed to forecast time series [14–17]. Statistics mining, connection mining, decision tree mining, NNM, and other techniques are used in data mining. ANN is an auto adapted dynamics system, which is one of these approaches. ANNs have developed as a potent statistical modelling technique and a highly powerful instrument in modern quantitative economics.

Various economic forecasting methods are presented to anticipate future growth of related industries based on existing state of industrial economy in terms of research methodologies. To forecast the tourism business, an FWA-SVR forecasting model was suggested. Grey model GM (1, 1)

was used by [18] to forecast the growth tendencies of China's real estate economy. Another study looked at China's macroeconomic scheme by using a BMF-VAR method to forecast the country's macroeconomy [19]. To predict and analyse China inflation rate, deep learning based LSTM method was presented in [20]. The author of [21] built a MIDAS model for assessing and forecasting CPI and PPI using commodity price big data. The Chinese consumer confidence index was analysed and forecasted using a DNN_CNN-LSTM model in [22]. Paper [23] suggested an autoregressive (AR) model for more accurate forecasting of energy demand and supply. Paper [24] created a dataset of 77 countries covering over 90% of world GDP and used cross-country statistics to anticipate US economic growth during downturns. In [25], the author offered a machine learning technique for constructing mood indicators as well as economic forecasting using evolving confidence indicators. Paper [26] developed a self-correcting intelligent pandemic prediction model to improve COVID-19 forecast accuracy and strengthen economic management and control. Paper [27] used WEKA3.9 to create three prediction models to accurately predict business failure.

The FWA-SVM model has the smallest errors in predicting the GDP ratio and filling rate [28] while it keeps the shortest time to optimize the SVM parameters. The bandwidth requirement of LSTM [29] is very high and the computational accuracy is low in these systems in big data analytics so BVNN_BP is proposed to overcome these limitations.

The proposed contribution of this paper is as follows:

(i) To collect the database and process it with proposed Bayesian vector neural network (BVNN) and to integrate with backpropagation (BP) model by feature extraction and classification

(ii) To compare the proposed BVNN_BP model with the existing system like FWA-SVR and LSTM with various datasets like WEO, APDREO, and AFRREO

However, understanding and applying these strategies necessitates a thorough understanding of mathematics, algorithm knowledge, and computer programme technology. Methodologies in the study case produce one type of simple and high-accuracy economic growth forecast.

2. Prediction Model Using Proposed Design of Bayesian Vector Neural Network (BVNN) with Backpropagation (BP)

2.1. Model Structure. Our BVNN_BP inputs are GDP growth index, primary, secondary, and tertiary industry growth value for n consecutive years in a region as panel data into the network to predict the output primary, and secondary and tertiary industry growth values for the region in year $n+1$ with attribute discretization, reduction importance ranking, and prediction rule.

The model divides input data into two parts.

(1) Regional GDP network: regional GDP data are independent of other indicator data before fully

connected layer of the neural network, and feature information is extracted using BP, batch normalisation, and ReLU activation function. This design reflects that the regional GDP can mirror the overall economic level of the region, which is equivalent to gauging the regional economic development level by regional GDP to divide regions and using different forecasting strategies for regions with different development levels in the fully connected layer.

(2) Regional panel data network: using neural network convolution kernels, the regional GDP growth index, primary, secondary, and tertiary industry growth values are integrated to produce panel data to detect industry correlations and effect of industry data on later growth values within n successive years. To extract time series data features as well as probable industry links, three BVNN layers are utilised first. (1) and (2) outputs values are inputs of fully connected layers and to enhance two networks multichannel outputs, GDP outputs are combined. The mathematical representation of BVNN layers is represented in Table1 with its algorithm.

2.2. Proposed Model Analysis.

Let us consider each feature by a variable $X_i \subseteq RR$. Then each client is totally defined by an n-dimensional vector. $x = (x_1 \ldots, x_n) \in X \subseteq R^n$.

Basic rule for calculating a unit's activation value in relation to other connected units with a strength w_{ji} is a function of weighted sum of inputs expressed [30] in

$$u_j = f\left(\sum_i u_i \cdot w_{ji}\right) = f(\text{Net}_j) = \frac{1}{1 + e^{-\text{Net}_j}}. \quad (1)$$

Net_j = Net Input to j level unit.

To this equation, one must include the unit's threshold, which is defined as the unit's proclivity to activate or inhibit itself by

$$u_j = f(\text{Net}_j) = \frac{1}{1 + e} - \left(\sum(i) w_{ji} \cdot u_i + \theta_j\right). \quad (2)$$

If $\partial u_j / \partial \text{Net}_j = u_j \cdot (1 - u_j)$ then Δout_j from (3) and (4):

$$\Delta \text{out}_j = (t_j - u_j) \cdot u_j \cdot (1 - u_j),$$

$$\frac{\Delta w_{ji}}{\partial} = -\frac{\partial E_p}{\partial w_{ji}}. \quad (3)$$

$$\begin{aligned}
E_p &= \frac{1}{2} \sum_k (t_{pk} - u_{pk})^2 \\
&= \frac{1}{2} \sum_k (t_{pk} - f_k(1 - u_j)) \\
&= \frac{1}{2} \sum_k \left(t_{pk} - f_k\left(\sum_i w_{kj} \cdot u_{pj} + \theta_k\right)\right).
\end{aligned} \quad (4)$$

E = Error; p = Model; t_k = Target; u_k = Output and then

$$\begin{aligned}
\frac{\partial(\text{Net}_{pk})}{\partial w_{kj}} &= \left(\frac{\partial}{\partial w_{kj}} \cdot \sum_j w_{kj} \cdot u_{pj} + \theta_k\right) \\
&= u_{pj} = -\frac{\partial E_p}{\partial w_{kj}} (t_{pk} - u_{pk}) \cdot f_k'(\text{Net}_{pk}) \cdot u_{pj}.
\end{aligned} \quad (5)$$

Value out_j will now determine the amount of value to be added or subtracted from weight w_{ji} in connection to activation state of unit u_i, which is activation with which u_j is linked to weight w_{ji} and coefficient r. This is correction rate that should be used as shown in

$$\Delta w_{ji} = r \cdot \Delta \text{out}_j \cdot u_j. \quad (6)$$

Δw_{ji} can have a negative or positive value. It denotes the "quantum" that will be added or subtracted from the prior weight w_{ji} value by

$$w_{ji(n+1)} = w_{ji(n)} + \Delta w_{ji}. \quad (7)$$

However, this assumption is only true for weights which connect a unit layer to output units layer. Correction approach outlined so far is only applied to BP with two layers, to use gradient descent [31] in (8) and (9).

$$\begin{aligned}
\frac{\partial J}{\partial w_{11}^1} &= \left(\frac{\partial J}{\partial a_1^2}\right)\left(\frac{\partial a_i^2}{\partial z}\right)\left(\frac{\partial z}{\partial w_1^1}\right) = \left(\frac{\partial\left((1/2)\left(a_1^2 - y\right)^2\right)}{\partial a_1^2}\right)\left(\frac{\partial(g(z))}{\partial z}\right)\left(\frac{\partial\left(a_1^1 w_{11}^1 + b^1\right)}{\partial w_{11}^1}\right) \\
&= \left(\left(a_1^2 - y\right)\right)(g(z)(1 - g(z)))\left(a_1^1\right) \\
&= (g(z) - y) \cdot g(z)(1 - g(z)) \cdot a_1^1.
\end{aligned} \quad (8)$$

It can be easily shown using calculus that derivative of the sigmoid function $g(z)$ is given by $g(z) \cdot (1 - g(z))$. Partial derivative of J is with respect to bias.

TABLE 1: Comparative analysis of regional economic prediction model based parameters for APDREO dataset.

Parameters	APDREO dataset		
	FWA-SVR	LSTM	BVNN_BP
GDP per capita prediction	88	90	92
Prediction accuracy	92	95	97
GDP per primary growth	90	93	95
GDP per secondary growth	89	92	96
Convergence rate	91	93	97

$$\frac{\partial J}{\partial b^1} = \left(\frac{\partial J}{\partial a_1^2}\right)\left(\frac{\partial a_1^2}{\partial z}\right)\left(\frac{\partial z}{\partial b^1}\right)$$

$$= \left(\frac{\partial\left((1/2)\left(a_1^2 - y\right)^2\right)}{\partial a_1^2}\right)\left(\frac{\partial(g(z))}{\partial z}\right)\left(\frac{\partial\left(a_1^1 w_1^1 + b^1\right)}{\partial b^1}\right)$$

$$= \left(\left(a_1^2 - y\right)\right)(g(z)(1 - g(z)))$$

$$= (g(z) - y) \cdot g(z)(1 - g(z)). \tag{9}$$

In this network, (10) can be rewritten as

$$\frac{\partial J}{\partial w_{11}^2} = \left(\frac{\partial J}{\partial a_i^2}\right)\left(\frac{\partial a_i^2}{\partial z}\right)\left(\frac{\partial z}{\partial w_{11}}\right)$$

$$= \left(\frac{\partial\left((1/2)\left(a_1^3 - y\right)^2\right)}{\partial a_1^3}\right)\left(\frac{\partial\left(g\left(z^2\right)\right)}{\partial z^2}\right)\left(\frac{\partial\left(a_1^2 w_{11}^2 + b^2\right)}{\partial w_{11}^2}\right)$$

$$= \left(\left(a_1^3 - y\right)\right)\left(g\left(z^2\right)\left(1 - g\left(z^2\right)\right)\right)\left(a_1^2\right)$$

$$= \left(g\left(z^2\right) - y\right) \cdot g\left(z^2\right)\left(1 - g\left(z^2\right)\right) \cdot a_1^2 \tag{10}$$

So, calculating the required gradients can be done. This explains the basic mathematics behind backpropagation. Now, after understanding the above mathematical aspects, we can backpropagate a simple linear NN using this approach. For more complex NN, with multiple nodes in every layer, the approach remains the same, and the only difference we get is the changing of subscripts and superscripts of w, z, a, and b. This makes backpropagation complex, and we need to use matrix forms of the derivatives, i.e., gradient vectors. But the basic idea remains the same. Backpropagation is simply using the chain rule of derivatives to calculate the change of cost function J with respect to changing parameters of the network. The proposed BP algorithm helps to compute the gradient which can be used to tweak the parameters using gradient as shown in Algorithm 1. Difference between actual output and desired one and summation report are constructed using previously calculated error coefficient Δout_j and weight associated with that coefficient was referring [32] to

$$\Delta\text{hidden}_i = u_i \cdot (1 - u_i) \cdot \sum \Delta\text{out}_j \cdot w_{ji}. \tag{11}$$

And therefore,

$$\Delta w_{ik} = r \cdot \Delta\text{hidden}_i \cdot u_k. \tag{12}$$

In fact, start from

$$\frac{\partial E_p}{\partial w_{kj}} = \frac{1}{2}\sum_k \frac{\partial}{\partial w_{kj}} \cdot \left(t_{pk} - u_{pk}\right)$$

$$= -\sum_k \left(t_{pk} - u_{pk}\right) \cdot f_k\left(Net_{pk}\right) \cdot w_{kj} \cdot f_j\left(Net_{pj}\right) \cdot u_{pi}. \tag{13}$$

w_{kj} = Hidden-Outputs weights.

Weights associated with output are corrected in the calculation as shown in

$$\Delta w_{ji} = r \cdot \Delta\text{out}_j \cdot u_i. \tag{14}$$

Weights that are not connected to output are corrected by

$$\Delta\text{hidden}_i = f'(u_i) \cdot \sum_j \Delta\text{out}_j \cdot w_{ji}. \tag{15}$$

Fulfilment of the corrections on weights is done by

$$w_{ij(n+1)} = w_{ij(n)} + \Delta\text{out}_i \cdot u_j \cdot \text{Rate}. \tag{16}$$

For output units as shown

$$\text{Bias}_{j(n+1)} = \text{Bias}_{j(n)} + \Delta\text{hidden}_j \cdot \text{Rate}. \tag{17}$$

$x_{t,i}^{sum}(\theta) \in R^n$ for all $t = 1, 2, \ldots, T$, where T is simulation length; i denotes random number initialized. In general, estimate or calibration processes try to find suitable values for h so that $X^{sim}(\theta, T, i)$ creates dynamics that are as near to those observed in an experimentally measured counterpart as possible as shown in

$$X = [x_1, x_2, \ldots, x_T]. \tag{18}$$

$x_t \in R^n$ for all $t = 1, 2, \ldots, T$.

As a result, Bayesian evaluation is phrased in terms of Bayes' theorem as given in

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)}. \tag{19}$$

Method for approximating $p(\theta|X)$ for a certain value of h then allows us to calculate right-hand side of

```
Algorithm for BP
Input:
h_i: current step size
τ: prescribed tolerance
ΔN_{i+1}, ΔN_i: Delta nodes Vectors
Local variables:
M_1: maximum factor by which step size is improved
M_2: factor by which step size will reduce if tolerance is not met.
q: safety factor so step size is just less than extreme suggested
𝒮←max|ΔN_{i+1} − ΔN_i|/h_i
if(h_i^2 𝒮/2) < τ then
h←min(q√(2τ/𝒮), h_iM_1)tolOK== true
Else
h←min(q√(2τ/ε), h_iM_2)
tolOK==false
end if
return ((tolOK, h̄ )h̄ ).
α_0 = (β_{1,...,r;0}, γ_0)': = α_0 = (β_{1,...,r;0}, γ_0)': = randomly initialized
η_0: = 1, t: = 0
do
α_old: = α_t
for each training examplexcalculateh(x; α_t)andz_j = f_1(x_i; β_t)
γ_{t+1}: = γ_t − η_t(h(x_i; α_t)−y_i)g_2'(γ_t'z_i)z_i, β_{j: t+1}: = β_{j: t} − η_t(h(x_i; α_t)−y_i)g_2'(γ_t'z_i)g_1'(β_{j,t}x_i)x_i., t: = t+1, η_t: = 1/twhile‖α_t − α_{ot d}‖ > ε
```

ALGORITHM 1: Proposed BP Algorithm.

$$p(\theta|X) \propto p(\theta|X)p(\theta). \qquad (20)$$

For all $t \geq \widetilde{T}$ in which $x_{t,i}^{sim}(\theta)$ varies for stationary level, $E[x_{t,i}^{sim}(\theta)|t \geq T]$ which allows us to further consider that $x_{\widetilde{T},i}^{sim}(\theta), x_{\widetilde{T}+1}^{\sin}(\theta), \ldots, x_{T,i}^{sim}(\theta)$ establishes a random sample from an assumed distribution. Using kernel density estimation (KDE), we can then establish a density function which designates this distribution, which we designate by $\widetilde{f}(x|\theta)$ permitting us to estimated likelihood of empirically sampled data for a specified value of h as given in

$$p(X|\theta) = \prod_{t=1}^{T} \widetilde{f}(x_t|\theta). \qquad (21)$$

For all $L < t \leq T$, this means that $x_{t,i}^{sim}(\theta)$ is solely dependent on past L realised values. As a result, our objective is to evaluate above conditional densities by

$$\widetilde{f}(x_{t-L,i}^{sim}, \ldots, x_{t-1,i}^{sim}, x_{t,i}^{sim}, \phi) \approx p(x_{t,i}^{sim} \mid x_{t-L,i}^{sim}, \ldots, x_{t-1,i}^{sim}: \theta). \qquad (22)$$

For all $L < t \leq T$, $\phi = \phi(\theta)$.

The aforementioned technique is made up of two main components: a mixture of K Gaussian random variables in Algorithm 2 and a random number generator as shown in

$$p(X|\theta) = \prod_{t=1}^{T-L} \widetilde{f}(x_t, \ldots, x_{i+L-1}, x_{i+L}, \phi). \qquad (23)$$

This leads us to define following loss function as shown in

$$LS(\theta^{rrue}, \widehat{\theta}) = \theta^{true} − \widehat{\theta}_2. \qquad (24)$$

3. Simulation Results

3.1. Dataset Description

3.1.1. World Economic Outlook (WEO).
The WEO database is developed during the biennial WEO exercise, which takes place between January and June each year and culminates in the release of the WEO in April and September/October. A database format is provided for selected series from the publication.

3.1.2. Asia and Pacific Regional Economic Outlook (APDREO).
APDREO offers data on recent economic trends and forecasts for Asian and Pacific countries. The data for REO for Asia and the Pacific is compiled in tandem with semiannual WEO exercises, which take place in the spring and fall. The data is in line with the WEO's projections. Due to variances in group membership, REO aggregate data may differ from WEO aggregate data. Weighted averages of data for specific countries make up composite data for country groups. All concepts are represented by arithmetic weighted averages, with the exception of inflation and broad money, which are represented by geometric averages.

3.1.3. Sub-Saharan Africa Regional Economic Outlook (AFRREO).
AFRREO provides data on recent economic events and scenarios in Sub-Saharan African countries. The data for REO for Sub-Saharan Africa is compiled in tandem with semiannual WEO exercises, which take place in the spring and fall.

```
Algorithm for BVNN:
Input:
F: finish time
N: number of particles
{R_0 ..., R_K}summary times
Local Variables:
S: vector of particles of sizeN
W: vector of weights of sizeN
Q: vector of summarised particles
h : proposed step sizeh ←R_1 − R_0
t_0←0; i←0; k←0
whileR_{k+1} < Fdo
forp = 1 ⟶ Ndo
i←0, t_i←R_kwhilet_i ≤ R_{k+1}do
t_{i+1}←t_i + hrepeat
S_p←samplefromP(x_{i+1}|x_i)
((tolOK, h )h )←Check Tolerance(h ).
if tolOK then
t_{i+2}←t_{i+1} + h. else
t_{i+1}←t_i + h. end if
until tolOK
i←i + 1end while
end for
W←F(y_{i+1}|x_{i+1})Q←summarise S . based onW.
S←. resample . based onW
k←k + 1
end while
return Q
```

ALGORITHM 2: Proposed BVNN Algorithm.

3.2. Comparative Analysis. This section discusses the comparative analysis for various datasets in terms of GDP per capita prediction, prediction accuracy, GDP per primary growth, GDP per secondary growth, and convergence rate. The techniques compared here are FWA-SVR and LSTM for WEO dataset, APDREO dataset, and AFRREO dataset. Python is the software used for implementation and the comparison results are obtained from the simulation. Table 2 represents the comparative analysis of regional economic prediction model based parameters for WEO dataset.

Table 2 and Figures 1 and 2 show comparative analysis for proposed BVNN_BP with existing technique LSTM and FWA-SVR. The parameters are compared for WEO dataset with existing techniques. Based on this comparison, the proposed technique obtained optimal results in terms of predicting the economic rate of the region and in terms of accuracy for this prediction higher for proposed technique when compared with existing technique. GDP predicted per capita by BVNN_BP is 97%, and accuracy is 98% which is higher based on the existing technique comparison. In terms of primary and secondary growth of GDP, it is comparatively higher for proposed model. Convergence is also optimized by BVNN_BP up to 91%. Table 1 shows the comparative analysis of regional economic prediction model based parameters for APDREO dataset.

Table 1 and Figures 3 and 4 show comparative analysis for proposed BVNN_BP with existing technique LSTM and FWA-SVR for APDREO dataset. This dataset gives the economic analysis and data from Asia and Pacific region.

For this regional analysis the GDP prediction obtained by proposed BVNN_BP is 92%, accuracy predicted by this proposed technique is 97%, and based on primary and secondary growth of GDP for this region of 95% and 96%, the convergence rate here is 97%.

Table 3 represents the comparative analysis of regional economic prediction model based parameters for AFRREO dataset.

Table 3 and Figures 5 and 6 show comparative analysis for proposed technique and existing technique for AFR-REO dataset. This dataset covers the economic growth for Sub-Saharan Africa region. Economically Africa is considered to be poor country, so the GDP prediction based on both primary and secondary industry growth is essential for this region. By analysing the data of this region, the GDP prediction is optimal but not up to the mark to region analysis. The accuracy for this dataset by proposed model would be 98%.

4. Discussion

Country econometric models continue to play an important role in planning and forecasting key indicators of socio-economic development. They allow taking into account the interdependencies of indicators associated with their economic content. At the same time, many economic time series are characterized by nonlinear dependencies, and the influencing factors are often impossible to describe explicitly

TABLE 2: Comparative analysis of regional economic prediction model based parameters for WEO dataset.

	WEO dataset		
Parameters	FWA-SVR	LSTM	BVNN_BP
GDP per capita prediction	94	95	97
Prediction accuracy	96	97	98
GDP per primary growth	85	87	92
GDP per secondary growth	89	93	95
Convergence rate	85	88	91
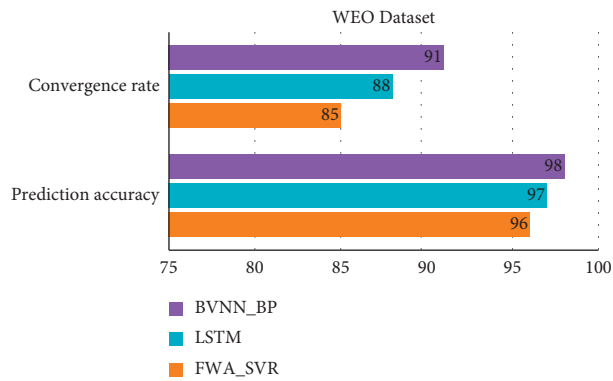
FIGURE 1: Comparative analysis of WEO dataset for convergence rate and prediction accuracy.
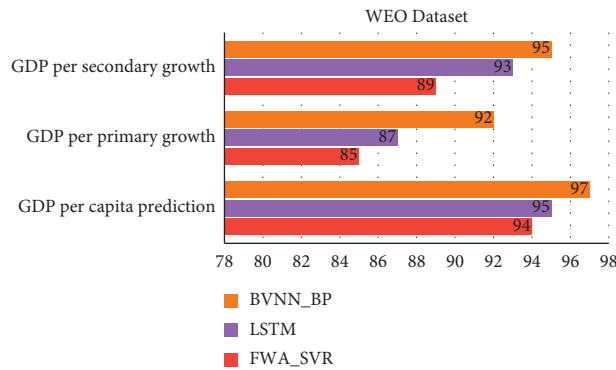
FIGURE 2: Comparative analysis of WEO dataset for regional economic prediction model.
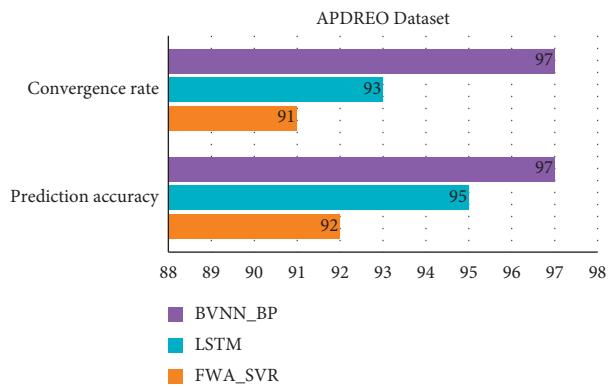
FIGURE 3: Comparative analysis of APDREO dataset for convergence rate and prediction accuracy.
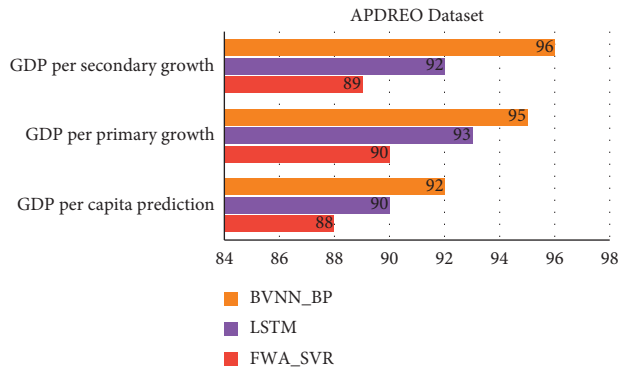
FIGURE 4: Comparative analysis of APDREO dataset for regional economic prediction model.

TABLE 3: Comparative analysis of regional economic prediction model based parameters for AFRREO dataset.

AFRREO dataset			
Parameters	FWA-SVR	LSTM	BVNN_BP
GDP per capita prediction	79	82	86
Prediction accuracy	92	95	98
GDP per primary growth	81	83	87
GDP per secondary growth	83	86	91
Convergence rate	85	89	93
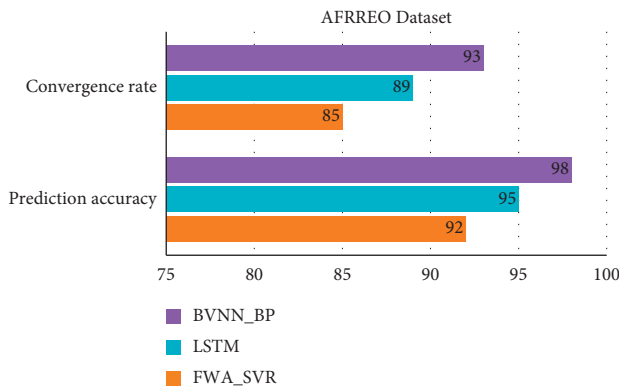
FIGURE 5: Comparative analysis of APDREO dataset for convergence rate and prediction accuracy.

in the form of regression equations. This makes, as shown by numerous studies, the application of neural network methods and technologies very promising.

Over the past decade, approaches to forecasting time series based on feedforward networks and recurrent networks have been developing, including their application for forecasting economic indicators. At the same time, machine learning platforms and tools are becoming widespread, making it possible to implement all the main architectures and methods of training neural networks. It should be noted that the use of neural network tools for separate unrelated time series does not allow for a systematic approach to modelling the economic sphere. We are convinced that a hybrid methodology should be applied with the possibility of choosing in each case the most appropriate method. Such an approach is implemented in developed Bayesian vector
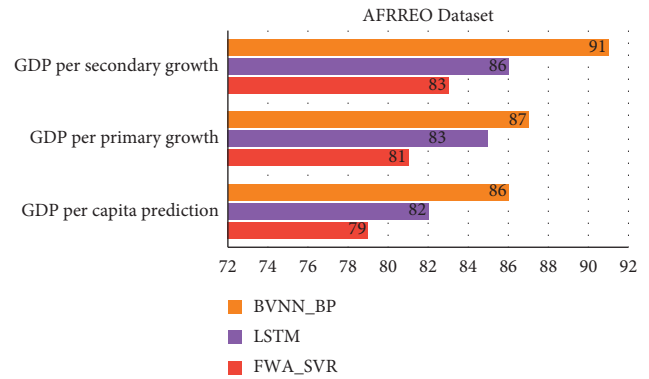
FIGURE 6: Comparative analysis of AFRREO dataset for regional economic prediction model.

neural network (BVNN) which is integrated with back-propagation (BP) based feature extraction and classification. The suggested technique can forecast regional economy and increase prediction model's convergence rate without impacting accuracy, effectively enhance effectiveness of regional economic analysis and prediction, lower percentage error, and have advantages of greater operating effectiveness. The only limitation of the proposed model is that the convergence rate and accuracy are dependent on each other so separate works are needed to improve the accuracy and convergence rate.

5. Conclusion

This paper proposed novel technique regional economic prediction based on machine learning techniques. Data were first normalised to balance deviations among indicators and then normalised data were transferred exponentially to balance deviations among regions, with the goal of reducing data heterogeneity among economic indicators of each region and differences in economic development levels of each region. Database has been collected and processed with proposed Bayesian vector neural network (BVNN) which is integrated with backpropagation (BP) model by feature extraction and classification. The experimental results show the comparative analysis for various datasets in terms of GDP per capita prediction, prediction accuracy, GDP per primary growth, GDP per secondary growth, and convergence rate. The techniques compared here are FWA-SVR and LSTM for WEO dataset, APDREO dataset, and AFRREO dataset. The proposed model reduces error by increasing the convergence rate and accuracy for each dataset.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

References

[1] S. Tuo, T. Chen, H. He et al., "A regional industrial economic forecasting model based on a deep convolutional neural network and big data," *Sustainability*, vol. 13, no. 22, Article ID 12789, 2021.

[2] Y. Dai and L. Su, "Economic structure analysis based on neural network and bionic algorithm," *Complexity Problems Handled by Advanced Computer Simulation Technology in Smart Cities*, vol. 2021, Article ID 9925823, 11 pages, 2021.

[3] H. B. Gu and M. Zhang, "Index analysis of regional economic integration based on internal forces and external forces: a case study of delta urban agglomeration near the Yangtze River," *Journal of Renmin University of China*, vol. 31, pp. 71–79, 2017.

[4] S. Garg, *A Study on the Structure of Neural Networks and the Mathematics behind Backpropagation*, 2019.

[5] D. Platt, "Bayesian estimation of economic simulation models using neural networks," *Computational Economics*, pp. 1–52, 2021.

[6] A. Chc, A. Fs, B. Fjh, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, 2020.

[7] J.-Y. Yeh and C.-H. Chen, "A machine learning approach to predict the success of crowdfunding fintech project," *Journal of Enterprise Information Management*, 2020.

[8] Z. Lv, Y. Li, H. Feng, and H. Lv, "Deep learning for security in digital twins of cooperative intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[9] C. G. Enright, M. G. Madden, and N. Madden, "Bayesian networks for mathematical models: techniques for automatic construction and efficient inference," *International Journal of Approximate Reasoning*, vol. 54, no. 2, pp. 323–342, 2013.

[10] S. Tuo and H. He, "A study of multiregional economic correlation analysis based on big data-taking the regional economy of cities in shaanxi Province, China, as an example," *Sustainability*, vol. 13, no. 9, 2021.

[11] G. Khatwani and P. R. Srivastava, "Impact of information technology on information search channel selection for consumers," *Journal of Organizational and End User Computing*, vol. 30, no. 3, pp. 63–80, 2018.

[12] T. Grubljesic, P. S. Coelho, and J. Jaklic, "The shift to socio-organizational drivers of business intelligence and analytics acceptance," *Journal of Organizational and End User Computing*, vol. 31, no. 2, pp. 37–64, 2019.

[13] L. Wu, Q. Zhang, C.-H. Chen, K. Guo, and D. Wang, "Deep learning techniques for community detection in social networks," *IEEE Access*, vol. 8, Article ID 96016, 2020.

[14] T. Yang, H. Pan, X. Zhang, A. Greenlee, and B. Deal, "How neighborhood conditions and policy incentives affect relocation outcomes of households from low-income neighborhoods—evidence from intra-city movement trajectories," *Cities*, vol. 119, 2021.

[15] L. Kong, Z. Liu, and J. Wu, "A systematic review of big data-based urban sustainability research: state-of-the-science and future directions," *Journal of Cleaner Production*, vol. 273, Article ID 123142, 2020.

[16] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," *Cities*, vol. 89, pp. 80–91, 2019.

[17] Z. Tang, M. Wang, D. Wei, and Z. Liu, "Seasonally-adjusted FWA-SVR model and its application in tourism economic forecast," *Journal of Systems Science and Mathematical Sciences*, vol. 41, 2021.

[18] Y. Ji, "Big data-based mixed frequency macroeconomic prediction and monitoring index construction," *Statistics & Decisions*, vol. 7, pp. 36–39, 2021.

[19] M. Xia and L. Jiang, "China's consumer confidence index forecast based on deep network CNN-LSTM model," *Statistics & Decisions*, vol. 7, pp. 21–26, 2021.

[20] P. Rakpho and W. Yamaka, "The forecasting power of economic policy uncertainty for energy demand and supply," *Energy Reports*, vol. 7, pp. 338–343, 2021.

[21] Y. Lyu, J. Nie, and S.-K. X. Yang, "Forecasting US economic growth in downturns using cross-country data," *Economics Letters*, vol. 198, Article ID 109668, 2021.

[22] O. Claveria, E. Monte, and S. Torra, "Economic forecasting with evolved confidence indicators," *Economic Modelling*, vol. 93, pp. 576–585, 2020.

[23] X. Tang, Z. Li, X. Hu, Z. Xu, and L. Peng, "Self-correcting error-based prediction model for the COVID-19 pandemic and analysis of economic impacts," *Sustainable Cities and Society*, vol. 74, Article ID 103219, 2021.

[24] S. Y. Kim and A. Upneja, "Majority voting ensemble with a decision trees for business failure prediction during economic downturns," *Journal of Innovation & Knowledge*, vol. 6, no. 2, pp. 112–123, 2021.

[25] Y. Huang, H. Wang, H. Liu, and S. Liu, "Elman neural network optimized by firefly algorithm for forecasting China's carbon dioxide emissions," *Systems Science & Control Engineering*, vol. 7, no. 2, pp. 8–15, 2019.

[26] W. Di, M. Wang, X. Sun, G. Han, and H. Xing, "Identification of bolt anchorage defects based on Elman neural network optimised by improved chicken swarm optimisation algorithm," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 62, no. 10, pp. 588–597, 2020.

[27] J. Yang, J. Wen, B. Jiang, and H. Wang, "Blockchain-based sharing and tamper-proof framework of big data networking," *IEEE Network*, vol. 34, no. 4, pp. 62–67, 2020.

[28] J. Cai, L. Yang, C. Zeng, and Y. Chen, "Integrated approach for ball mill load forecasting based on improved EWT, refined composite multi-scale dispersion entropy and fireworks algorithm optimized SVM," *Advances in Mechanical Engineering*, vol. 13, no. 2, Article ID 1687814021991264, 2021.

[29] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: convolutional LSTM model for sentiment analysis in social big data," *Information Processing & Management*, vol. 58, no. 1, Article ID 102435, 2021.

[30] A. Kumar, R. Shankar, and N. R. Aljohani, "A big data driven framework for demand-driven forecasting with effects of marketing-mix variables," *Industrial Marketing Management*, vol. 90, pp. 493–507, 2020.

[31] R. D. Raut, S. K. Mangla, V. S. Narwane, B. B. Gardas, P. Priyadarshinee, and B. E. Narkhede, "Linking big data analytics and operational sustainability practices for

sustainable business management," *Journal of Cleaner Production*, vol. 224, pp. 10–24, 2019.

[32] M. Ghayekhloo, R. Azimi, M. Ghofrani, M. B. Menhaj, and E. Shekari, "A combination approach based on a novel data clustering method and Bayesian recurrent neural network for day-ahead price forecasting of electricity markets," *Electric Power Systems Research*, vol. 168, pp. 184–199, 2019.

Hindawi

Research Article

Recognition of Bookmark Aging Degree Based on Probabilistic Neural Network

Cong Zheng, Xiaoling Zhang, Shaoqiu Ma, and Zhijian Xiao (ID)

The School of Mathematical Engineering, Zhejiang Dongfang Polytechnic, Wenzhou 325000, China

Correspondence should be addressed to Zhijian Xiao; xiaozhijian77@stu.ahu.edu.cn

Bookmarks are the basis for librarians to get books on and off shelves and borrowers to borrow books. In order to solve the problem of time-consuming and labor-consuming manual checking of bookmark aging, this paper proposes a method of bookmark aging recognition based on image processing technology. First, we perform image preprocessing, Otsu threshold segmentation, and morphological processing on the acquired bookmark image to obtain the effective area of the bookmark, then acquire the aging features for the bookmark, and finally input the acquired features into the trained neural network for defect recognition. The experimental results show that the method proposed in this paper can achieve 96% recognition, which can more accurately identify the aging defects of bookmarks.

1. Introduction

With the proposal and continuous innovation of the concept of smart library [1], the trend of intelligent digitalization of libraries has become increasingly obvious. Different from digital library [2], smart library emphasizes the intelligent management of library. Besides the management of digital resources, it is necessary to develop an intelligent system for paper book management, so as to enhance the state detection and ensure paper books are in good condition. At present, automatic borrowing and returning of books has been realized, that is, readers can not only check the borrowing status and information of the required books through the mobile phone app or the computer provided by the library but also complete the borrowing and returning procedures in the self-service book borrowing and returning machine [3].

Libraries often check books through inventory, so that the problems of books, such as misplaced books [4], damaged books, and bookmark falloff, can be found [5]. For example, the school library of the author is closed for arrangement every Friday morning. For libraries with large collections, this method will not only affect the normal borrowing of books but also pose great challenges to

librarians. In 2016, more than 100 librarians in Nanjing University Library spent 5 days to complete the inventory of above 900,000 books [6]. In view of the shortcomings of traditional library management, such as being time-consuming and labor-consuming, having low efficiency, and common complaints of "unable to find books," some libraries have developed library robots, such as the intelligent book inventory robot in Nanjing University Library, the RFID book automatic inventory robot in Wuhan University Library [7], and the intelligent chat robot Xiaotu in Tsinghua University Library [8]. Library robot has the advantages of saving manpower, noncontact, time management, and efficiency, which can solve the problems of traditional library management to a certain extent. Particularly, in terms of manpower, time, and cost, it can fundamentally improve the library management and make routine management possible. However, according to the current studies and reports, some library robots can only realize the handling in designated positions, such as the handling robots of Humboldt University Library in Germany and Osaka City University Library in Japan [9]. Other library robots need to customize bookshelves or use other devices to keep books in the designated state, such as the automatic access robots of Johns Hopkins University in the United States [10] and National University of Singapore [11].

The workload and cost of bookshelf update and modification are massive. Although the current robot technology has been relatively mature, it can hardly be popularized in library management. Therefore, for the current large libraries, it is more appropriate to apply robot technology to library management based on the current actual situation of libraries than to completely adopt new technology. According to the existing studies of intelligent book management, the combination of robots and other technologies is more suitable. In this study, the book state detection algorithm is mainly developed based on image processing technology, which can be superimposed on the robot. The robot is equipped with a camera to realize the image collection. The proposed algorithm is used to detect and judge book images and provide corresponding suggestions to assist the daily management and maintenance. Different book detection algorithms will be continuously developed for later robot implantation, so as to effectively reduce the times of internal arrangement by closing library and decrease the workload of librarians in library management.

2. Book Detection Framework

As a bridge among books, readers, and librarians, bookmarks play a vital role in book borrowing and book shelving. Figure 1 shows several common types of the bookmark.

Different from the two-dimensional code of books, the bookmark is usually attached to the lower part of the spine. It is easily exposed to the air and affected by the frequent circulation of books, resulting in aging, damage, and even falling off. Some common problems of bookmarks are shown in Figure 2.

A aging bookmark detection algorithm based on image processing is proposed, whose overall flow is shown in Figure 3.

Firstly, image segmentation is carried out for the obtained bookmark image, the bookmark area is segmented, and five features of aging classification are extracted from the segmented bookmark area, which are gray mean, median and variance of B channel, and the mean and median of S channel. The extracted features are input to the trained neural network model, and three aging states of bookmarks are obtained, namely, non-aging, slight aging, and severe aging.

3. Detection of Bookmark Aging

According to the bookmark images, bookmarks usually have the following characteristics:

(a) Bookmarks are composed of the white background with black characters, and there are edges of other colors with a certain width.

(b) When bookmarks are aged, they will become yellow or even brown. The more serious the aging, the darker the color.

Based on the above-mentioned characteristics of bookmarks, a bookmark aging detection method based on visible light images is proposed. With the robot as the carrier, a high-definition pan/tilt camera is used to shoot bookmark images. First, the bookmark area is extracted by image preprocessing, segmentation, and extraction algorithm. Second, according to the color characteristics, the aging characteristics of bookmarks are determined. Finally, the aging degree of bookmarks is judged by inputting the bookmark characteristics to the trained aging detection model. The flowchart is shown in Figure 4.

3.1. Bookmark Segmentation. Before aging detection, it is necessary to extract the label from the image, and the integrity of the label region extraction is the premise to ensure the aging recognition accuracy. Therefore, before detecting the aging of bookmarks, this study first determines the bookmark area. The main methods include image preprocessing, image segmentation, and morphological processing. Figure 5 shows the process of label region extraction.

(a) *Image Preprocessing.* Preprocessing of bookmark images [12] is not only the premise of bookmark aging but also an important step of bookmark recognition and aging degree detection. The main purpose is to make bookmarks obvious in images, so as to separate bookmarks and remove the background information. Firstly, the images are grayed and converted from RGB images to gray images, which is conductive to simplifying subsequent detection process and decreasing computation burden.

(b) *Image Segmentation and Morphological Processing.* After preprocessing the collected bookmark images, the bookmark area is preliminarily extracted by image segmentation [13], and the segmented images are subjected to morphological processing [14] to remove the bookmark's noisy point. Meanwhile, bookmarks are separated from the background according to the unique distribution characteristics, thus determining the bookmark area.

Because of the special structural features of bookmarks, a unified brightness map is obtained through camera imaging, and the light reflection ability is stronger than that of the background. After image preprocessing and image segmentation, an accurate bookmark area can be obtained, which provides a guarantee for subsequent processing.

3.2. Aging Degree Detection. Aging is inevitable in the use process, but the causes are complex. Aging problem cannot be solved fundamentally. Therefore, the best solution is to regularly check and then replace aged bookmarks. In this study, neural network is used to classify the aging degree. Before detection, relevant characteristic parameters are extracted, and bookmarks are detected by using PNN (probabilistic neural network). The detection mode is shown in Figure 6.

3.2.1. Feature Extraction. Because of obvious color changes caused by bookmark aging, the recognition of the aging degree can be transformed into color recognition. The
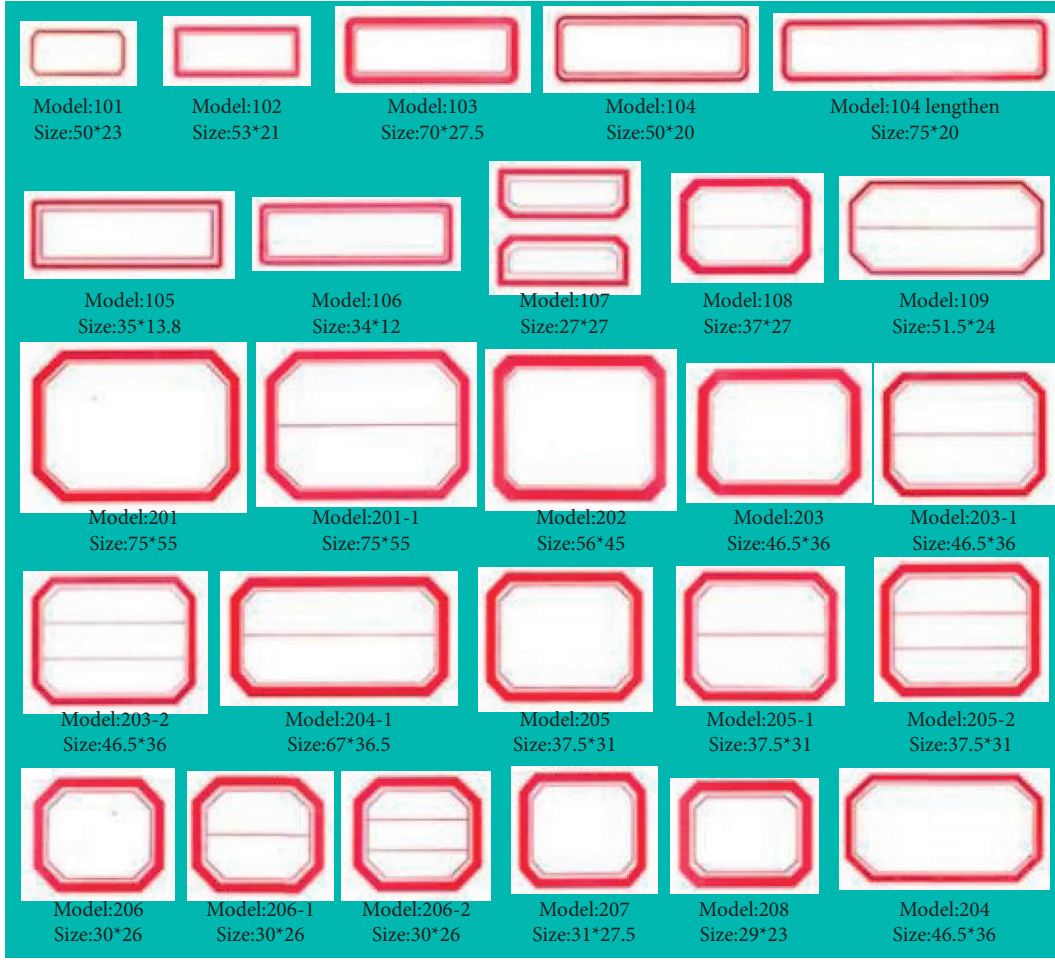
FIGURE 1: Bookmark types.

gray values of R, G, and B channels [15] will decrease during aging, and the change of B channel is the most obvious. HSI color space [16] is composed of hue (H), saturation (S), and intensity (I) separately. S component is less affected by external factors, and its performance is relatively stable. Therefore, this study selects B and S components to extract feature values [17] (including mean value, median value, and variance) of the two channels (B and S) in the bookmark area, respectively (five features, namely, B_{mea}, B_{var}, B_{med}, S_{mea}, and S_{med}). The feature calculation formula is as follows.

Mean value:

$$G_{mea} = \sum_{i=1}^{m} \sum_{j=1}^{n} g_{i,j} p_{i,j}. \tag{1}$$

Median value:

$$G_{med} = \underset{1 \le i \le m, 1 \le j \le n}{\text{median}} g_{i,i}. \tag{2}$$

Variance:

$$G_{var} = \sum_{i} \left(g_{i,j} - G_{mea} \right)^2 p_{i,j}, \tag{3}$$

where $g_{i,j}$ represents the gray values in the B and S images; $p_{i,j}$ is the probability of gray value $g_{i,j}$; m and n represent the image size; and $g_{i,j}$ is the operation of taking the median value.

According to the field test, three groups of sample data of bookmark aging degree can be obtained, as shown in Table 1. The second column of Table 2 refers to the three aged bookmarks obtained from the collected image (the aging degree is non-aging, slight aging, and severe aging), and the corresponding five feature values are listed in the third to eighth columns, respectively. It can be found that the extracted features have obvious differences in different aging degrees, and it is effective to select the five features for classification.

3.2.2. Aging Detection Model. PNN (probabilistic neural network) [18] is a feed-forward artificial neural network put forward by Dr. Specht in 1990 on the basis of Bayes rule and radial basis neural network. It is composed of radial basis neuron [19] and competition neurons [20]. As shown in Figure 7, the model is mainly composed of four layers: input layer, pattern layer (training set), summation layer, and output layer. It can use linear learning algorithm to complete
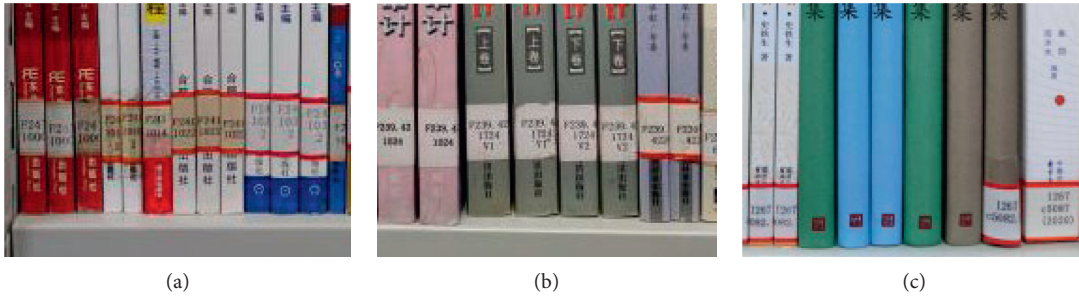
Figure 2: The common problems of bookmarks. (a) The problem of aging. (b) The problem of damage. (c) The problem of falloff.
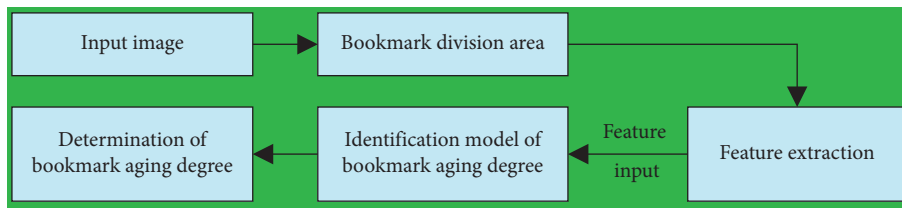
Figure 3: The overall flow of the bookmark detection algorithm.
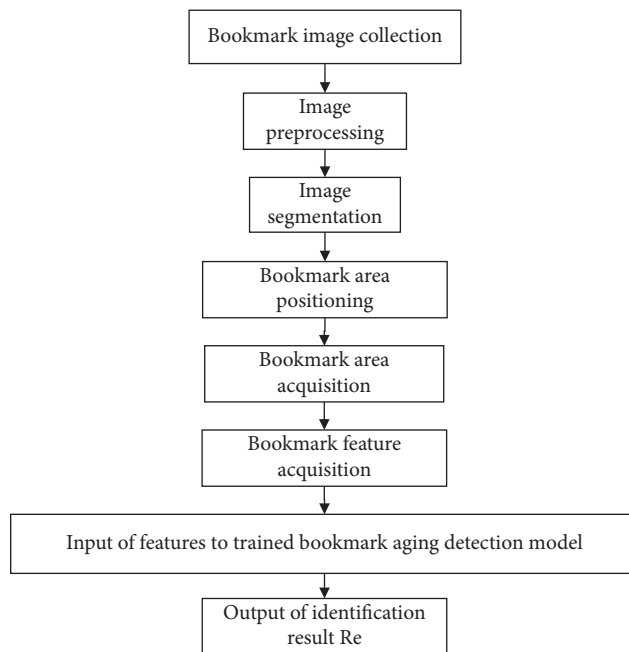
Figure 4: The bookmark aging detection algorithm.
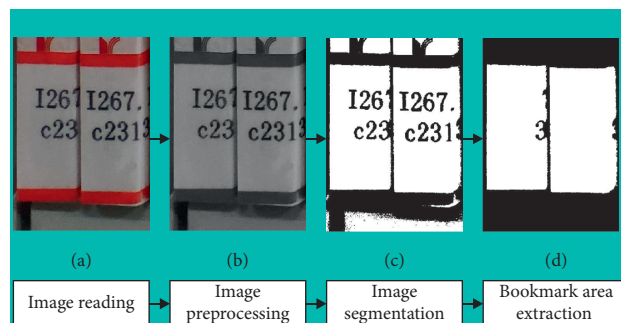
Figure 5: The process of the bookmark area extraction.
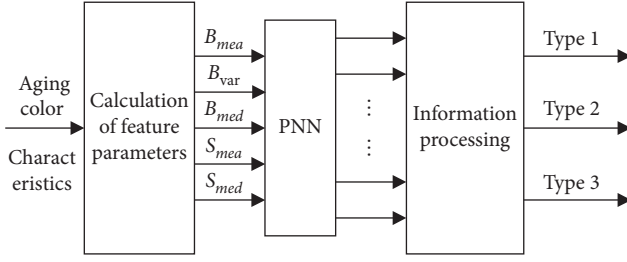
Figure 6: The model of the bookmark aging detection.

the function of non-linear algorithm while maintaining the high precision of non-linear algorithm. With the characteristics of simple structure and quick training, the model is suitable for pattern classification.

(1) Classification of Probabilistic Neural Networks.

Input layer: the input layer receives data and transmits them to the mode layer. The number of input neurons is the same as the dimension of input vector.

Pattern layer: the pattern layer calculates the matching between input vector and each pattern of the training set. The number of training samples determines the number of neurons in the pattern layer. Neurons correspond to training samples. Each pattern layer calculates the output result according to the following formula:

$$\phi_{ij}(x) = \frac{1}{(2\pi)^{d/2}\sigma^d}\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right], \quad (4)$$

where σ is the smoothing coefficient; x is the input vector composed of calculated color features; and x_{ij} is the neuron vector of the i_{th} pattern layer of class j.

Summation layer: neurons in the summation layer correspond to the classification patterns, which are connected with the pattern layer through sparse links. The connection condition of neurons between the summation layer and the pattern layer is that they correspond to the same classification pattern. By summing the output values, the maximum possibility of a certain category can be obtained. The calculation formula is shown in the following formula:

$$p_j(x) = \frac{1}{(2\pi)^{d/2}\sigma^d}\frac{1}{N_j}\sum_{i=1}^{N_j}\exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right], \quad (5)$$

where N_j is the number of samples of class j.

Output layer: the class with the highest posterior probability is selected as the classification result according to Bayesian decision rules, and the calculation formula is as follows:

$$C(x) = \arg\max\{p_j(x)\}, \quad j = 1, 2, \ldots, m, \quad (6)$$

where $C(x)$ is the estimation class of the input vector x and m is the total number of classes.

(2) Construction and Training of Network. There are five nodes in the input layer of PNN (probabilistic neural network), which correspond to five characteristic parameters (B_{mea}, B_{var}, B_{med}, S_{mea}, and S_{med}). There are three nodes in the output layer, which correspond to three states, namely, non-aging, slight aging, and severe aging. These three states are expressed in binary format, as shown in Table 2.

100, 010, and 001 represent the categories of non-aging, slight aging, and severe aging bookmarks output by the network, respectively. In the subsequent testing process, in order to draw the test result diagram, the description corresponding to the aging category will be converted into decimal representation, that is, 1, 2, and 3 represent 100, 010, and 001, respectively.

4. Algorithm Experiment

In order to verify the effectiveness of the proposed bookmark aging detection method based on the PNN (probabilistic neural network), experiments were carried out on a large number of collected bookmark images, including image segmentation, neural network training and testing, and aging degree detection. The experiment was completed on a computer with Intel Core i5-3240 @3.40 GHz CPU and 4G memory and image processing software MATLAB R2014a.

4.1. Bookmark Segmentation Effect. The bookmark segmentation proposed in this paper is mainly based on maximum between-class variance (Otus method) threshold segmentation and morphological filtering. The segmented bookmark area is mapped back to the original image to obtain the bookmark area without background interference. The processing results are shown in Table 3. The second to sixth lines represent the original image, the grayed image, the threshold segmented image, the morphologically processed image, the bookmark area determined by the minimum circumscribed rectangle, and the color image of the bookmark area.

According to the Otsu segmentation results in Table 3, it can be found that the bookmarks aged in different degrees are correctly segmented, and the whole area is complete without false segmentation. The proposed segmentation method is effective and feasible for the bookmark segmentation.

4.2. Aging Degree Identification Test. In this study, the bookmark images taken in the experimental environment are selected. 60 images of non-aging, slight aging, and severe aging samples are taken as the training samples, respectively. The input layer of the neural network is (B_{mea}, B_{var}, B_{med}, S_{mea}, S_{med}), and the output layer has three aging degrees (nonaging, slight aging, and severe aging). B_{mea}, B_{var}, B_{med}, S_{mea}, and S_{med} of the training sample are extracted as the input of the neural network, and the aging degree is the output for PNN neural network training. 150 images of samples with different aging degrees (50 images of non-aging, slight aging, and severe aging in each category) are selected as the test samples to simulate the neural network, and the corresponding aging degree

TABLE 1: The data of the aging bookmark.
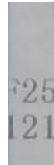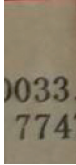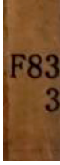
Serial number	Original image	Type	B_{mea}	B_{var}	B_{med}	S_{mea}	S_{med}
1		Non-aging	181.22	8.13	188	0.0372	0.0416
2		Slight aging	107.96	6.78	115	0.2453	0.2447
3		Severe aging	46.17	2.94	47	0.6077	0.6226

TABLE 2: The classification of aging modes.

Aging classification	Description of binary format	Description of decimal format
Non-aging	100	1
Slight aging	010	2
Severe aging	001	3
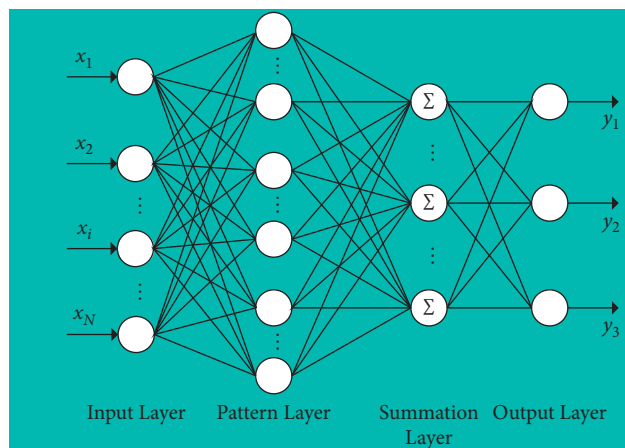
FIGURE 7: Probabilistic neural network structure diagram.

discrimination results are output. There are obvious differences among bookmarks with different aging degrees in terms of B_{mea}, B_{var}, B_{med}, S_{mea}, and S_{med}. The actual output and the target output of 150 test samples are shown in Figure 8, in which Yc-Label is the experimental test result and T-Label is the actual test result.

As can be seen from Figure 8, all slight aging bookmarks are correctly identified, and three non-aging bookmarks and severe aging bookmarks are mistakenly identified as slight aging. In pairwise transition, the values of the extracted features may be similar, and the typical samples fail to cover adequately, resulting in recognition errors. Such problems can be solved by accumulating the number of samples.

In this study, the backpropagation (BP) neural network and the competition network are also simulated and tested. As shown in Table 4, the accuracy of the neural network is 96%, which is obviously higher than that of the BP neural network [21] and the competition network, thus realizing the accurate judgment of the gray density degree of bookmark aging.

This study selects five bookmark images to analyze the gray aging degree of bookmarks, which are compared with the field manual test results. The results are shown in Table 5. The proposed method can detect and judge the actual aging degree of bookmarks, and the results have good consistency.

TABLE 3: The process of the image segmentation.

Serial number	1	2	3	4	5	6	7	8	9	10
Original images										
Image preprocessing										
Image segmentation										
Morphological processing										
Bookmark area extraction										
Color images										
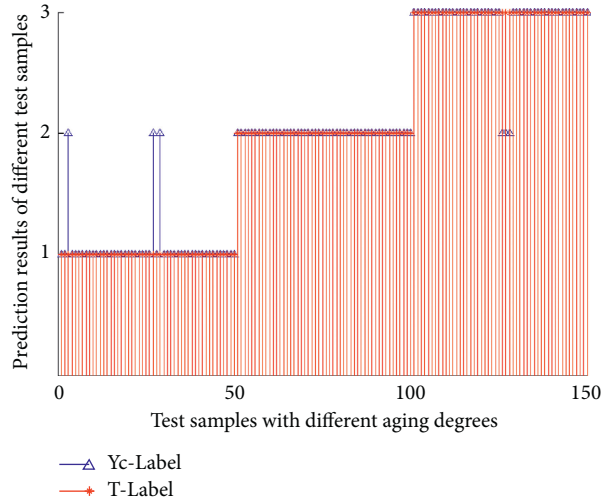
FIGURE 8: Comparison of PNN's detection and actual situation.

TABLE 4: The result of aging degree judgment.

Aging degree	Actual number	Correctly judged number		
		Competition network	BP	PNN
Non-aging	50	50	44	47
Slight aging	50	37	45	50
Severe aging	50	41	45	47
Total	150	128	134	144
Recognition rate	—	85.33%	89.33%	96%

TABLE 5: The process of image recognition.

Serial number	1	2	3	4	5	6	7	8	9	10
Original images										
Color images										
Segmented bookmark area										
Bookmark aging detection area										
Recognition results										
Actual results	3	3	1	1	2	1	2	2	1	1	3	3	3	1	2	2	2	2
Consistent or not	Consistent	Consistent	Consistent	Consistent	Consistent	Consistent	Consistent	Consistent	Consistent	Consistent

5. Conclusions

The bookmark defect detection method based on probabilistic neural network can make full use of the prior knowledge of faults and qualitatively detect the aging of bookmarks under Bayesian minimum risk criterion. Probabilistic neural network has high training speed. It is easy to implement in engineering and has strong robustness and high recognition accuracy. With the gradual accumulation of fault knowledge, the network can be continuously expanded to further improve the detection accuracy.

This study proposed the identification of bookmark aging degree based on probabilistic neural network and applied it to detect the aging degree of bookmarks. The identification accuracy is 96%, which is higher than that of the traditional method. The proposed method lays a foundation for the detection of bookmark defects by using image processing technology.

However, the proposed method only analyzes the aging of bookmarks, but there are many other defects in bookmarks. Therefore, the algorithm will be improved to further study other types of bookmark defects.

Data Availability

The characterization data used to support the results and discussion are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Zhijian Xiao was responsible for sorting out data, data collection, and literature review. Cong Zheng and Shaoqiu Ma participated in the experiment and data analysis. Xiaoling Zhang was responsible for full-text writing.

Acknowledgments

References

[1] Z. Wu and X. Yang, "Development of research and practice of smart library in China," *Library and information work*, vol. 65, no. 4, pp. 20–27, 2021.

[2] Y. Ning and Z. Zhang, *New Positioning And New Challenges Of Digital Library In The Era Of Scientific Big Data -- A Summary Of The 16th Advanced Seminar On Frontier Issues Of Digital Library. Books and Information*, vol. 40, no. 1, pp. 127–135, 2020.

[3] X. Ding, "Comparative analysis of Library self-service loan and return system," *Library and information work*, vol. 62, no. 22, pp. 78–81, 2018.

[4] l Jiang, Y. Jiao, B. Wang, and Z. Ding, "Design and research of library intelligent shelf error reminder system based on RFID technology," *Library theory and practice*, no. 5, pp. 27–30, 2019.

[5] H. Wang, "How to protect the book labels, bar codes and seals of collected books," *Library work and research*, no. 10, pp. 101-102, 2008.

[6] H. Fan and B. Shao, *Innovative Practice and Thinking of Intelligent Robot Book Inventory -- Taking Nanjing University Library as an Example. Library*, vol. 36, no. 9, pp. 96–100, 2018.

[7] Z. Xia, L. Quan, W. Duan, and P. Fu, "Application research of RFID book automatic inventory robot -- taking wuhan university library as an example," *Library Journal*, vol. 39, no. 1, pp. 61–66, 2020.

[8] L. Li, "reconstruction and innovative development of library service model from the perspective of artificial intelligence -- analysis based on the report of artificial intelligence: opportunities and impact of future decision-making," *Books and Information*, no. 6, pp. 30–36, 2017.

[9] J. Fang and Y. Zhang, "Research and application of automatic book access robot in Library," *Library construction*, no. 7, pp. 79–83, 2012.

[10] J. Suthakorn, S. Lee, Y. Zhou, R. Thomas, S. Choudhury, and G. Chirikjian, "A robotic library system for anOff-site shelving facility," in *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*, pp. 3589–3594, Washington, DC, USA, May 2002.

[11] K. H. Yuan, A. C. Hong, G. S. Peng, and M. Ang, "Unmanned library: anintelligent robotic books retrieval & return system utilizing RFIDTags," in *Proceedings of the 2002 IEEE Systems, Man and Cybernetics*, Yasmine Hammamet, Tunisia, October 2002.

[12] x Zhuang, C. Wang, and Z. Sun, "ACT image preprocessing method based on attention mechanism," *Small microcomputer system*, pp. 1–7, 2021, Available at: http://kns.cnki.net/kcms/detail/21.1106.TP.20210420.1103.015.html.

[13] Y. Zhu, Z. Ling, and Y. Zhang, "Research progress and Prospect of machine vision technology," *Journal of graphics*, vol. 41, no. 6, pp. 871–890, 2020.

[14] G. Li and T. Wang, "Research on concrete crack detection based on morphological processing and feature analysis," *Building structure*, vol. 50, no. S2, pp. 529–533, 2020.

[15] G. Jiang, J. Xiao, and Z. Huo, "Pedestrian re recognition method integrating RGB and gray image features," *Computer Engineering*, vol. 47, no. 4, pp. 226–233, 2021.

[16] Y. Zhou, Z. Ma, X. Shi, and T. Kong, "Surface defect detection method of linear guide rail in HSI color space," *China Mechanical Engineering*, vol. 30, no. 18, pp. 2179–2184, 2019.

[17] J. Sun, X. Ding, L. Du, Q. Li, and F. Zou, "Research progress of fabric image feature extraction and retrieval based on convolutional neural network," *Journal of textiles*, vol. 40, no. 12, pp. 146–151, 2019.

[18] Y. Zhang, Y. Jia, W. Wu, X. Su, and X. Shi, "Application of probabilistic neural network in typical fault diagnosis of vehicle gearbox," *Automotive Engineering*, vol. 42, no. 7, pp. 972–977, 2020.

[19] H. Tang, q. Li, and Z. Liu, "Research on meteorological data interpolation method based on rough RBF neural network," *Computer engineering and design*, vol. 35, no. 1, pp. 282–286, 2014.

[20] N. Feng, Y. Dong, S. Nan, and Z. Guo, "The method of balanced competition neuron winning chance," *Computer engineering and design*, vol. 30, no. 4, pp. 971–973, 2009.

[21] Y. Pan and L. Ma, "Gearbox fault diagnosis based on BP neural network," *Coal mining machinery*, vol. 42, no. 5, pp. 165–167, 2021.

Research Article

News and Public Opinion Multioutput IoT Intelligent Modeling and Popularity Big Data Analysis and Prediction

Hao Yan (ID)

School of Education and Modern Art, Shangqiu Institute of Technology, Shangqiu, Henan 476000, China

Correspondence should be addressed to Hao Yan; 1350008031@sqgxy.edu.cn

Based on the news and public opinion multioutput Internet of Things architecture, this article analyzes and predicts its popularity with big data. Firstly, the model adopts a three-tier architecture, in which the bottom layer is the data layer. It is mainly responsible for the collection of the terminal sensor data of the Internet of Things, and it uses intelligent big data as the data warehouse. Secondly, the computing layer on the data layer mainly provides the computing framework. Using the open-source SQL query engine, a cluster environment based on memory computing is constructed to realize the parallelization of data computing. It is used for interactive operations between the system and users. It receives and forwards the query requests submitted by the client browser, transmits them to the server cluster for execution, and displays the results in the browser. The end is displayed to the user. After that, combined with the needs of the development and application of news and public opinion big data, the data collection process was analyzed and designed, and the distributed data collection architecture was built. The intelligent Internet of Things was adopted for data storage, the data storage structure was analyzed and designed, and the data storage structure was designed to avoid catching. The repeat check algorithm is used to repeatedlystore the obtained page data. At the same time, according to the analysis of the business needs of the news and public opinion information platform, the overall functional structure of the platform was designed. The database and platform interface were designed in detail. The simulation results show that the model realizes the statistical query of the collected sensor alarm information and historical data on the user system, combines the historical operating data to analyze the relationship between the supply/return water temperature of the heat exchange station and the outdoor temperature, and realizes chart visualization of data analysis.

1. Introduction

With the development of modern information technology, emerging technologies, such as the Internet, cloud computing, and big data, have been widely used in various fields of social economy. The sensor transmission speed of the Internet of Things is very fast, and the application of a large number of sensor devices will definitely lead to a substantial increase in the output of news and public opinion data. The Internet of Things and big data are closely related, and the data generated by sensors can also be processed by the big data platform [1, 2]. News and public opinion Internet of Things big data is different from Internet big data [3]. In addition to the general characteristics of big data, they also have strong relevance and timing. Therefore, the traditional Internet big data processing methods are not fully

applicable. New solutions need to be designed specifically to properly analyze the Internet of Things data and extract more important information from the Internet of Things monitoring equipment [3–5].

To improve the efficiency of the analysis of news and public opinion IoT data, it is necessary to implement a low-latency query system based on the Hadoop data warehouse, which can handle a large number of concurrent query requests. A single query can return the results faster, thereby improving the efficiency of data processing and analysis [6]. With the increasing maturity of the wireless sensor network technology,the combination of RFID, other sensors and wireless communication technology can realize the information exchange between objects,as well as long-distance monitoring and management if connected to the Internet. The internet realizes human-to-human interaction with

people as the object, and the Internet of Things will further realize the interaction between things and realize the concept of ubiquitous networks. Using distributed query strategies and memory-based computing methods, the main purpose is to quickly query the information in massive amounts of data, provide timely feedback to users, and improve query efficiency. For massive data, traditional processing methods are difficult to meet real-time requirements. Distributed technology has become a research hotspot in the field of big data [7–9].

Based on this, the system takes the news and public opinion as the research object and uses computer network and communication technology under the framework of the Internet of Things to store the news and public opinion-related information and deliver it to the users in a timely and accurate manner, realizing news and public opinion detection, management, and marketing network. The traditional news and public opinion detection technology has been improved to improve the efficiency of news and public opinion detection, personnel management, and business decision-making. In the process of platform construction, a data storage architecture based on Mysql and intelligent Internet of Things was designed, SpringMVC development framework was adopted. Spring Framework was used as the core container, and Dubbo was used as the distributed architecture of the entire platform.The platform system is powerful, safe and scalable, which can handle high concurrency and massive data storage.Dubbo chooses intelligent big data as a data storage tool and creates news, public opinion, and IoT big data based on an open-source memory computing engine. Query the system and use data mining algorithms to analyze the historical data of the operation of the heat exchange station in the news and public opinion detection.

2. Related Work

With the continuous increase in the popularity of big data research, the application fields of big data are becoming more and more extensive. Domestic internet companies use big data technologies, such as Hadoop, to handle PB-level data problems in data storage, data mining, and high concurrency. Using Hadoop-based architecture for data collection, data analysis, etc., major universities are also conducting academic research and applications in the big data environment. Under the general situation, big data has been widely used in internet finance, medical health, transportation, and communication operations for other aspects [10–12].

Losavio et al. [13] proposed a news and public opinion big data acquisition and analysis platform (IBDP) that integrates HDFS, Spark, intelligent big data, HBase, Flum, Sqoop, OpenStack, etc., suitable for the acquisition and analysis of news and public opinion data. Yigitcanlar et al. [14] proposed and developed a smart city system based on the Internet of Things using the Hadoop ecosystem and big data analysis technology, combined with Spark over Hadoop to achieve the efficiency of big data processing. Hossein Motlagh et al. [15] proposed the use of the Hadoop software environment, including data collection, data storage, data normalization and analysis, and data visualization components to realize the

parallel processing of large heterogeneous data for IoT network security monitoring. Scholars have developed a farmland observation data management system based on the integration of wireless sensor networks that realize the automatic acquisition of a large amount of news and public opinion data. The system has been applied in the city for the related processing and analysis functions. Chin et al. [16] studied news and public opinion big data from the connotation of news and public opinion big data, the acquisition of news and public opinion big data, and the status quo of news and public opinion big data and combined the current internet technology and big data technology to look forward to news and public opinion big data.

Van Deursen and Mossberger [17] elaborated on the four key technologies of big data research: data collection and preprocessing, data storage, data analysis and mining, and data presentation and application, and gave the architecture diagram of big data collection, data warehouse, parallel storage architecture.In addition, they also introduced the high-availability technology of mass storage system, parallel computing, real-time computing, streaming computing, deep learning, data privacy protection technology, and other related technologies, and provided reference learning cases. According to the characteristics of massive data, the researcher designed the system structure of the massive data management system based on Hadoop and introduced the distributed storage and distributed computing of the massive data in detail, which has a strong reference value [18]. At present, there are many types of enterprise information management systems, and each system is managed independently, resulting in the wastage of resources and poor scalability of the system. In this backdrop, Lu designed a hierarchical storage system using the Hadoop key technology to design data mining function, data migration module, etc., to provide data mining and data migration system based on Hadoop architecture. It shows that with the continuous update of internet technology, the complexity of the news and public opinion management information system is also increasing. However, few news and public opinion management information systems can use all the functions of the internet and the emerging accurate news and public opinions and face the news and public opinion management information. The system's requirements for accurate news and public opinion and traditional management information systems require that the implementation of these systems is more complex. Some studies are based on the need of identifying accurate news and public opinions and realizing news and public opinion management by evaluating modern networks [19–21].

3. Based on News and Public Opinion Multioutput IoT Intelligent Modeling and Popularity Big Data Analysis and Prediction Model Construction

3.1. Intelligent IoT Hierarchical Nesting. At the intelligent IoT level, data analysis needs to move a subset of data to the data warehouse, and the speed of data analysis in

Hadoop is very slow. However, with the development of SQL query engines, big data technology can already be used in business analysis scenarios. By building a data model in Hadoop or other databases, large-scale historical data accumulated and stored for a long time are used in the big data processing system for information mining [22]. Figure 1 is the hierarchical topology of the intelligent Internet of Things.

The data collected by the sensors play an important role in detection monitoring and data mining analysis. In addition to real-time monitoring of data generated by IoT terminal devices, it is also necessary to store historical data accumulated in the process of news and public opinion detection and provide real-time statistical query analysis and the function of generating data reports.

$$L(m, t) = \frac{l(m, t)}{N},$$

$$\psi(x) = \frac{f(x, d) - f(x, m)}{f(x, d) + f(m, d)}. \tag{1}$$

In the data monitoring system of IoT devices, the sensor devices transmit the monitoring data to the data processing platform using various transmission methods, such as HTTP, TCP, and MQTT and store them in MySQL after parsing to provide real-time query support. The massive historical data adopts the intelligent big data storage warehouse to provide large-scale data support for the data mining and analysis of the system.

$$\chi(x, \lambda) = \begin{cases} e^{-\lambda L \times (1 + \lambda x^2)}, & \chi \geq \chi(x), \\ 0, & \chi < \chi(x), \end{cases}$$

$$L(y, g) = \sum_{i=1}^{n} y_i \left(g_i + \ln \times \sum_{j=1}^{n} \exp(g_i) \right). \tag{2}$$

When performing statistical query and data mining analysis on historical data, a query engine based on memory calculation is used to improve the speed of the system query and analysis. At the application layer, the output of the big data platform layer is used for chart display and to build a web server platform to provide a visual interface for data display and analysis. The processing of data collected by the terminal equipment includes real-time monitoring, statistical query analysis, and data mining analysis.

$$\begin{cases} f(x) = \text{sign} \left\{ \sum_{i=1}^{N} a_i y_i \times k(x_i, x_j) - b \right\}, \\ \text{sign}(x) = \begin{cases} -x > 0, \\ x = 0, \\ x \leq 0. \end{cases} \end{cases} \tag{3}$$

The terminal collection system consists of wireless IoT sensing devices, gateways, and data storage servers. The terminal devices regularly transmit data to the gateway through a network protocol. The collected data is parsed by the gateway and stored in a unified format.

3.2. Multioutput Process of News and Public Opinion.
News and public opinion have different output data nodes, and the collected data information will also have different degrees of difference. It is necessary to unify the data access methods of all IoT application terminals and storage standards for different forms of data at different nodes and unify the storage and management of multisource data. At the same time, it is necessary to provide hybrid computing capabilities for multisource data to improve the efficiency of multisource data management and analysis. The web server receives the request submitted by the client browser and sends the query task to the Presto computing cluster, and finally, it returns the result of the query execution to the user and displays the result on the client browser in the form of a web page for the data analyst to provide a friendly visual interface. The front-end is mainly implemented with Ajax, JavaScript, jsp, CSS, html, and other technologies, and ECharts is used to realize chart visualization. Figure 2 is a multioutput fan chart of news and public opinion.

When the coordinating node of the ZigBee network receives the single-point-sent data packet from itself or other devices, the network layer will continue to pass it on according to the program we set up in advance according to the requirements. If the target node is its neighbor, the data packet will be directly transmitted to the target node. If the target node is not its neighboring node, the coordinator node will retrieve the record that matches the destination address of the data to be transmitted and the routing table. If there is a match, the data will be transmitted to the next level network in the record; otherwise, the router will initiate a path search. When a node device receives an RREQ data packet, it will forward the data packet, in turn, and add the latest connection cost value. By analogy, the RREQ data packet follows all the connections it passes through and carries the sum of all connection costs until the RREQ data packet successfully reaches the destination device node. The routing node will select the one with the lowest connection cost among the received RREQ packets and send a route reply packet RREP to the source device. The RREP packet is a single-point sending packet that will return to the node device that sent the request along the path where the RREQ packet came from. After the path is established, the data packets can be sent and received along the established path. At this time, the node will send an RERR packet to all the node devices waiting to receive the data of this node to set the path as invalid. Each node will also update its routing table according to the received data packet.

3.3. Popularity Demand Analysis.
In the computing structure of popularity, the data stream is composed of countless tuples. It is the smallest unit of data that contains many key-value pair data. The spout is the entrance to the data source. It provides many simple API interfaces, including sensor output API interfaces, website hits, messages from social networking sites, and logs of various applications. It converts the received data source into a tuple data stream and transmits the data to the next component specified. The main purpose of writing a spout program is to obtain data from the data source. Bolt is a function in the storm
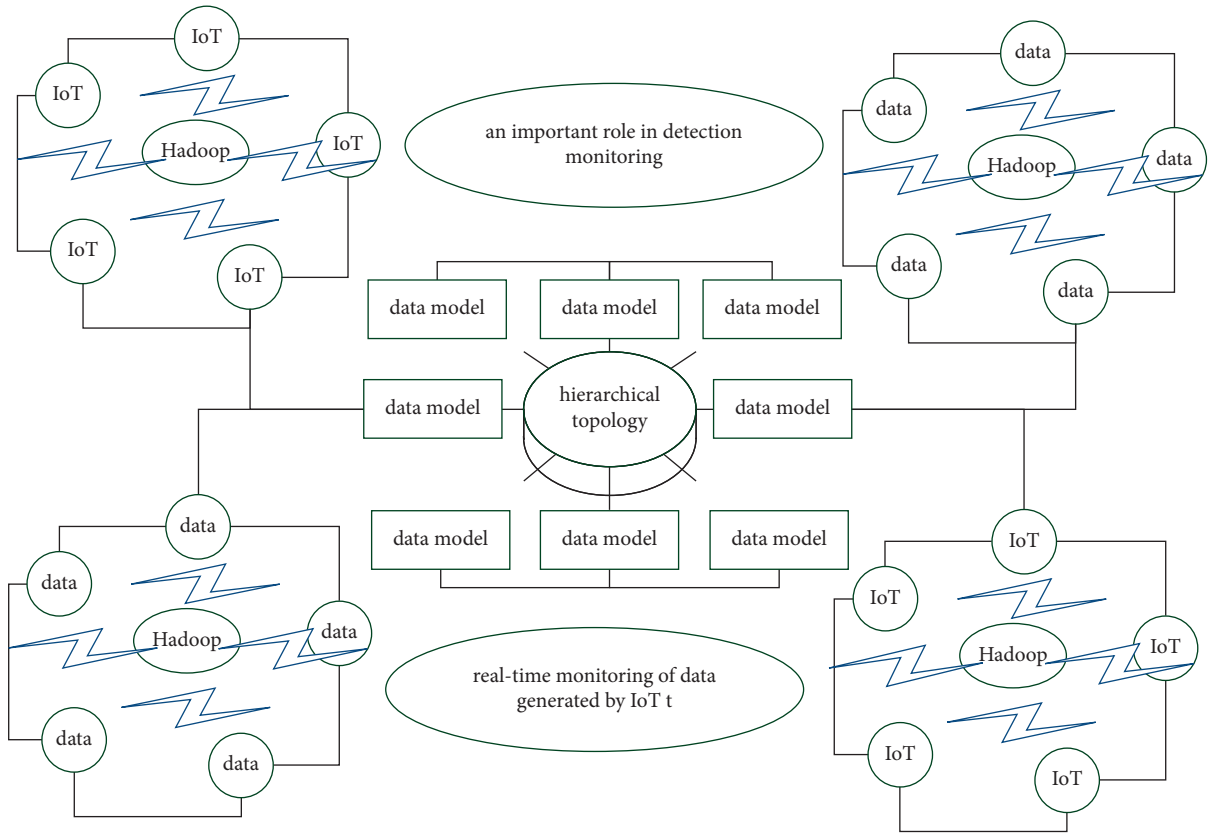
FIGURE 1: Smart IoT hierarchical topology.
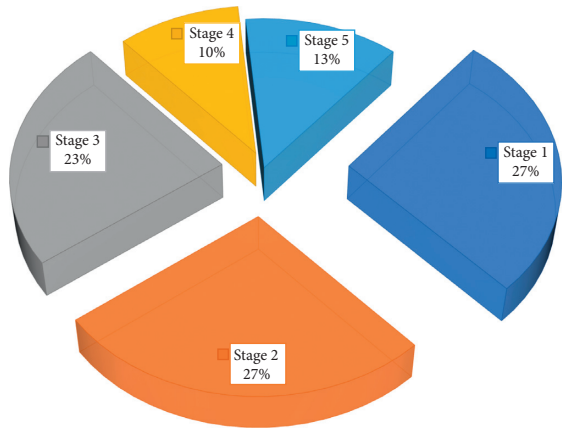
FIGURE 2: Multioutput fan chart of news and public opinion.

program. It is responsible for calculating and processing the input data stream. Its main functions include data filtering, data fusion, data calculation processing, and writing data to or reading data from the database. The terminal sensor device sends data to a unified data processing platform through the gateway so that each device can share and exchange data seamlessly. At the same time, it integrates the processing of real-time data and historical records and provides a unified operating environment for the application platform of the Internet of Things. Integrate demand analysis and solutions in a variety of application scenarios so that the system can meet the various needs of the users.

Figure 3 shows the popularity requirement scalability architecture. NewSQL mainly refers to the improved SQL database with scalability and superior performance. As it is an improvement and innovation based on the original SQL database, the NewSQL is compared to the original SQL database technology. Since it is an improvement and innovation of SQL technology, it still has the functions of traditional SQL, supports SQL queries, and meets the transactional and consistency requirements of the database queries. At the same time, the improved NewSQL database also has scalability and flexibility similar to NoSQL databases. HDFS supports the redundant backup storage of data blocks. As HDFS requires a combination of relatively low-cost small computers, these small computers are not highly reliable, and hence, HDFS is designed to be highly fault-tolerant. It is able to detect and respond to the failure of each machine node in time and ensure the stability of the system.

For this part of the data source, a variety of data acquisition interfaces need to be developed in the system, including the RS485 interface, Ethernet interface, AD conversion interface, and 24 V switch detection interface. By deploying data to HDFS, it can support large-capacity, high concurrency, and high-throughput big data computing tasks, while keeping file systems consistent across the nodes. An HDFS cluster must contain a name node NameNode (aka master node) and multiple DataNodes (aka slave nodes). Each slave node in the HDFS system is an ordinary or cheap computer. The name node can provide a naming service for each storage unit in the HDFS system, record and
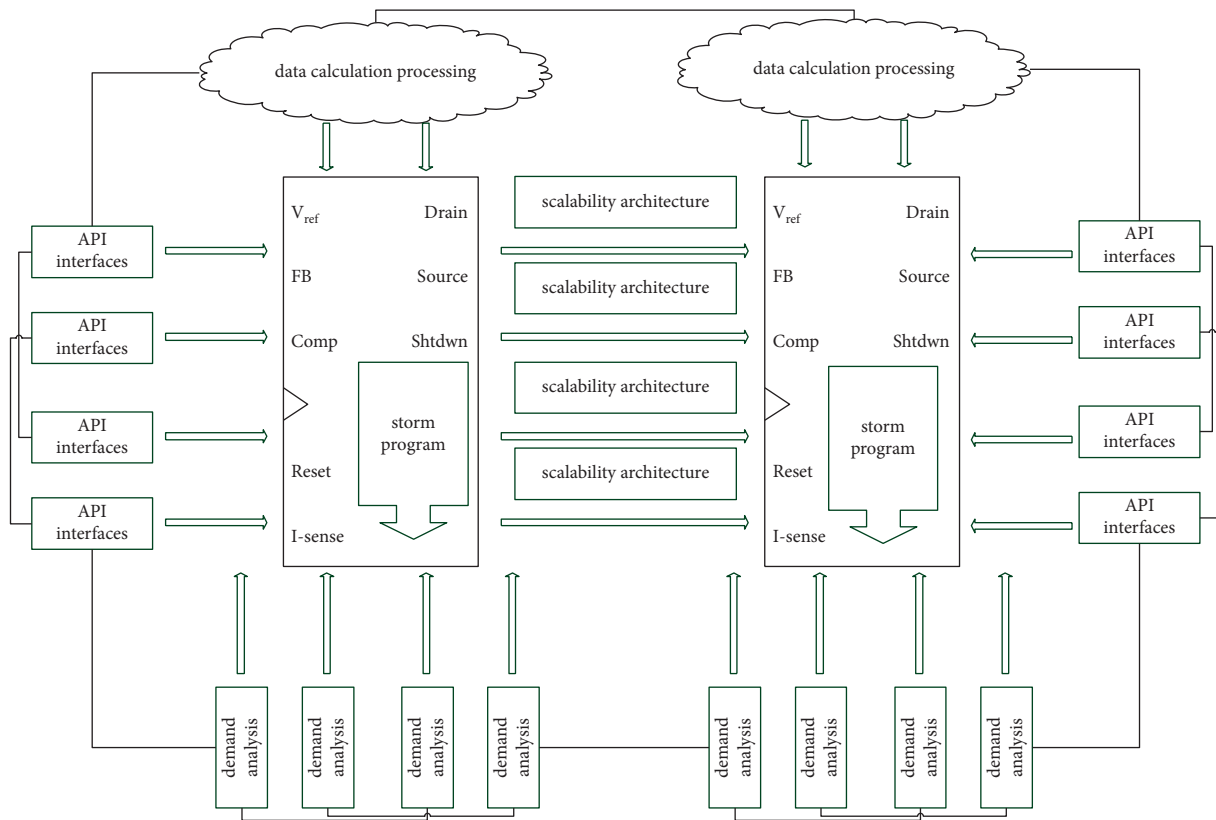
FIGURE 3: Popularity requires scalability architecture.

maintain the mapping information of the entire system data block, and receive requests to access HDFS from the corresponding client. Data nodes are mainly used to perform specific tasks, such as storing file blocks scheduled by the client and the NameNode. At the same time, HDFS has a client interface that interacts with the outside world. It mainly implements the external access requests to HDFS, including interacting with the NameNode to read file storage information, interacting with the DataNode to read the data in the HDFS system, and so on.

3.4. Design of Big Data Forecasting Function.

The Hadoop system has many components, including HDFS, MapReduce, HBase, smart big data, Sqoop, Flume, Zookeeper, etc. It can run efficiently on the Linux platform. It supports multiple programming languages and has high reliability and fault tolerance. It can process the data reliably and efficiently and has relatively good scalability. It is used to store and process large amounts of data for analysis, and thus, Hadoop has become a very popular solution. When the Hadoop cluster becomes insufficient, it is usually improved by adding new computers or storage devices with common methods. The Hadoop ecosystem can perform distributed computing well, and the users can develop without knowing the details of the underlying storage. The data collection terminal mainly relies on wireless network sensor equipment for real-time data collection. The wireless node performs data collection after registering in the network. At the

same time, the data is sent to the gateway through the MQTT protocol and Modbus bus at regular intervals. The data of multiple nodes are summarized. The collected data is preprocessed and sent to the relational database server in a unified format, and the application server side processes the data. Figure 4 is the distribution of prediction accuracy of public opinion big data.

In streaming computing, the data continuously flows into the system. The streaming computing system analyzes and calculates the continuous data in real-time and quickly in the memory, and then, it feeds the results back to the user in real-time or stores them for subsequent queries. Traditional streaming computing systems are mostly designed based on event mechanisms, and the amount of data that can be processed is limited. However, new streaming computing technologies, such as S4, Storm, and Spark, are mainly oriented to the use of streaming processing. The NameNode enables the clients to quickly access the required blocks for regular operations, adopts block placement strategies and replication mechanisms to ensure data availability and durability, and allocates new block locations while maintaining the load balance of the cluster. The main task of the DataNode is to store data. When a new data request is stored in HDFS, it will be split into blocks with a fixed preconfigured size, stored in the DataNode, replicated a fixed number of preconfigured times, and stored in different nodes. The DataNodes and NameNode generally communicate by a heartbeat message mechanism every few seconds so that the NameNode can know which node is unavailable
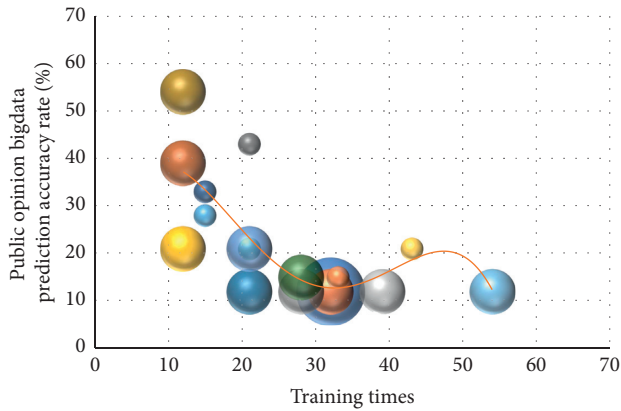
Figure 4: Distribution of prediction accuracy of public opinion big data.

Figure 5: Distribution of news and public opinion data transmission query speed.

and other useful information about the node. The client directly communicates with the DataNode when accessing data. When the client forwards the request to HDFS, the NameNode, firstly, sends back the location of the block required by the client after verifying the relevant license, and then, the client directly commands the DataNode to execute the required block.

4. Based on News and Public Opinion Multioutput IoT Intelligent Modeling and Popular Degree Big Data Analysis and Prediction Model Application

4.1. Smart IoT Big Data Query. This paper uses ATMEGA32-16AU type single-chip microcomputer of ATLEL Company, which is a low-power and high-performance 8 bit AVR microprocessing chip that uses an RISC structure. The single-chip core contains a rich instruction set and up to 32 general-purpose registers with 32K bytes of in-system programmable flash, 2K bytes of SRAM, 1K bytes of EEPROM, and 32 general-purpose I/O interfaces to meet the needs of the system. The I/O interface of ATMEGA32 one-chip computer can visit through the IN and OUT order and carry on data transmission between 32 general registers and I/O interfaces. The address of the register from 0x00 to 0xlf can be directly addressed by CBI and SBI commands, and the value of a certain bit in the address can be detected by SBIS and SBIC.

Figure 5 is the distribution of news and public opinion data transmission query speed. When querying, the embedded device sends a query command to the server. The server queries the database according to the command type and displays the data result on the embedded device via the coordinator, routing node, and terminal node. The entire architecture of the platform is mainly based on MySQL and MongoDB databases with Spring Framework as the core container combined with Zookeeper as the registration center and Apache Shiro as the authority authorization layer. My Batis is used as the persistence of the data access layer, and Redis is used as the cache database to improve the database access speed. In the data analysis layer, the Flume
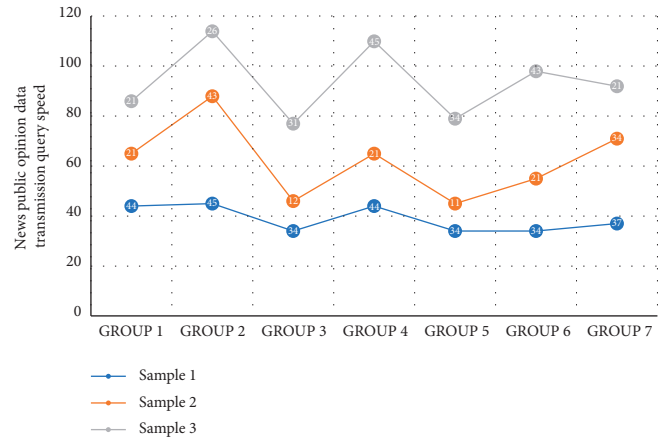
component can obtain a large amount of heterogeneous data in the HDFS storage system and supply it to the Hadoop offline batch processing system for analysis and processing. The processing results are written into the corresponding database. The Kafka component obtains the sensor network data stream and provides it to the Storm real-time data processing system for data analysis. The Storm system will process the analyzed data and write it into the corresponding database using the bolt component. Finally, there is the application layer. This layer queries and reads the data in the database according to the requirements of different applications or reads the processing and analysis results of the storm in real time.

4.2. Multioutput Platform Simulation of News and Public Opinion. When researching the news and public opinion information platform based on big data, the paper chooses the data storage mode combining Mysql and intelligent Internet of Things. Mysql is mainly used to store traditional business data, such as user information table, permission table, payment information table, agricultural material shop table, order information table, customer service information table, agricultural material evaluation information table, etc. However, the platform collects a large amount of user behavior data and a variety of types. Relational databases, such as Mysql can no longer meet the requirements. For alarm data and cloth length data with high real-time requirements, it is necessary to increase the frequency of data collection. The intelligent Internet of Things is stored in the BSON structure. Mass data storage has obvious advantages, and hence, the authors choose to use the intelligent Internet of Things to store the user behavior data. At the same time, for the business data on Mysql, a data backup is made in the intelligent Internet of Things. Among them, the data acquisition module uses various types of high-precision sensors and converts the collected analog signals into digital signals that can be identified and processed by the chip. In terms of communication designed in the thesis, the ZigBee communication technology is used for short-distance wireless transmission, and the GPRS and GSM

communication networks are used for long-distance communication. The processing control module paper uses a single-chip microcomputer for processing information, a relay for controlling operations, a memory chip, etc. An integrated circuit board is made. Figure 6 is the distribution of multiple output signals of news and public opinion.

As the big data system mostly uses NoSQL database technology, a comparative study of this type of database technology is carried out so that a suitable database system can be selected according to the needs. There are many classification methods for NoSQL data inventory, and various classification methods may also overlap. If the database is compared according to the four types of data models, namely the key-value model, column model, document model, and graphical model, and the specific correlations of the four types of data models are shown in Figure 6. Therefore, it can be concluded that the increase in the amount of calculation data does not weaken Storm's computing power, which indicates that the data are effectively cached under the action of the Kafka component and ensures the smooth and efficient operation of the storm system. Figure 7 is the distribution of news and public opinion calculation data.

When a device that has joined the network receives this sentence and the RejoinNetwork parameter is set to 0x00, NLME will send an NLMEJoIN. The confirm statement and the parameter value is INvALI. When a device that is not currently joining any network receives this sentence and the RejoinNetwork parameter is set to 0xol, the device will try to join the network specified by ExtendPANId, and then, NL dirty will issue MLME-AsSOCIATE. CoorAddress parameter setting is an address determined according to the situation of the router. This statement defines the initialization of the upper-layer device, allowing the upper-layer device to start a new ZigBee network and use itself as the coordinator. The Beaconorder parameter in the code represents the command for the network beacon formed by the upper layer. Super deorder means the command for the network superframe formed by the upper layer. BatteryLifeExtension means that if the value of this parameter is TRUE, NL will request coordination.Otherwise, NL will request that the coordinator does not support the battery life extension mode. The dynamic data sources of the warp knitting machine are diverse, including the warp let-off PLC of the warp knitting machine, electric meters, and various sensors. If a semifunctional node receives this sentence, NLME will send out an NLME-NETw0RK-FORMAT10 N whose parameter status is set to REQUEST confirm statement. The network cannot be established at this time. If a coordination node receives this sentence command, the device will be initialized as a coordination.

4.3. Example Application and Analysis.
Considering the above factors, this paper chooses the 10T-NODE2530 module. The sensor network node provided by this module is very complete. It supports the Zidie 2007 Pro wireless communication protocol very comprehensively. The node uses the CC2530 chip based on TI SoC. The chip's FLASH

capacity is 256I (B, the module uses a standard interface for expansion, which can be expanded according to different application requirements because HBase runs on Hadoop, and hence, one needs to build a Hadoop cluster first. The cluster consists of four machines. It consists of four nodes. Among them, Node1 is the NameNode node, and the other nodes are the DataNode nodes. The wireless module is connected with the server using the RS232 level interface, and the serial port converter is connected to the upper computer. The serial port converter adopts the six-in-one multifunction of Technology Co., Ltd. The serial port module CP2102, USB, TTL, RS232, and RS485 four levels can be switched using the switch to realize the six serial conversion functions of USB to TTL, USB to 232, USB to 485, TTL to 232, TTL to 485, and 232 to 485. Figure 8 shows the conversion efficiency distribution of big data wireless modules.

GPRS specifies four forms of channel coding, namely CS-1, CS-2, CS-3, and CS-4, and their corresponding data rates are 9.05 kbps, 13.4 kbps, 15.6 kbps, and 21.4 kbps. The transmission rate is proportional to the wireless environment. It can be seen that the Cs-4 channel coding method has the highest requirements for C/I. At present, GPRS has been developed to support the multicoding mode and multislot technology fusion transmission, and its maximum speed can reach 171 kbps. The system software structure consists of a data source, query engine, and application server. The specific structure of the data query framework is shown in the text, where each component can be multiple in the system. Using MySQL and smart big data as the connection data source, MySQL is mainly used as a real-time query database. Smart big data can be used as a data warehouse for storing historical data. In the cluster, the query plan is executed by Presto. Firstly, the data is converted from the serial port data into IP data, and then, it is sent out through the GPRS transmission system. The processing control module sends data to the GPRS module through the Rs232 serial port. The packet data will be encapsulated by an SGSN, and the data will communicate with the gateway support node GGSN via the GPRS backbone network after being encapsulated. Figure 9 is the distribution of news and public opinion serial port data acceptance rate.

In this experiment, the GPRS long-distance wireless transmission module is selected. The module uses the design scheme of the built-in protocol stack of the module. There is no chip outside the module, and hence, the module has higher stability. It supports up to 4 network connections and can connect to the serial port. The received data is sent to 4 servers at the same time. The module is set with a keep-alive mechanism to ensure that the network connection will not be disconnected when there is no heartbeat packet. It supports remote configuration parameters that can be configured by sending AT commands via SMS. The commonly used method of measuring the distance between the two nodes is based on the difference in the arrival time. This method uses the sending node to send two signals with different propagation speeds at the same time. Firstly, the receiving node calculates the difference between the arrival times of the two signals. Secondly, it combines their
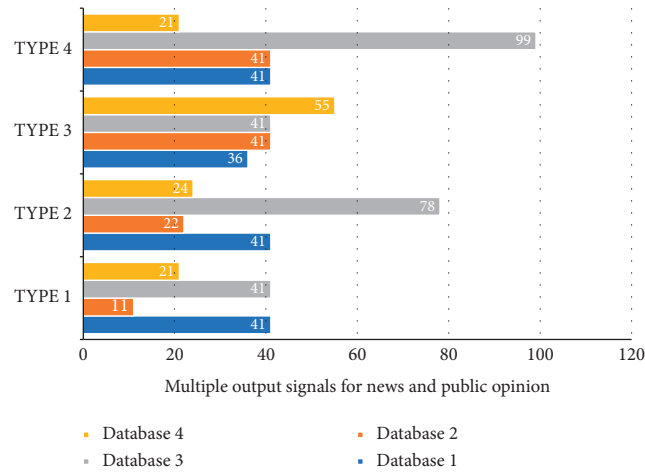
FIGURE 6: Multioutput signal distribution of news and public opinion.
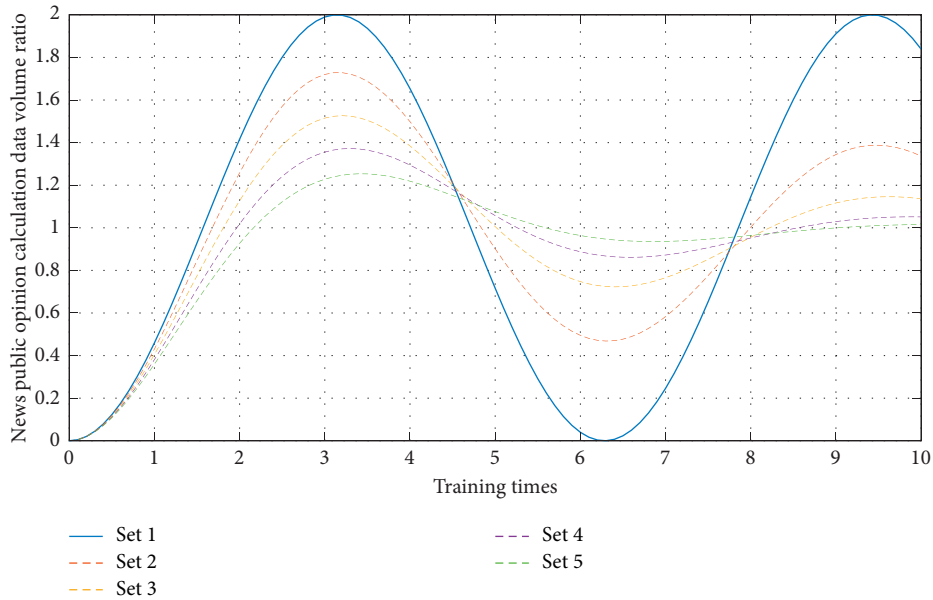
FIGURE 7: Distribution of data volume for news and public opinion calculation.
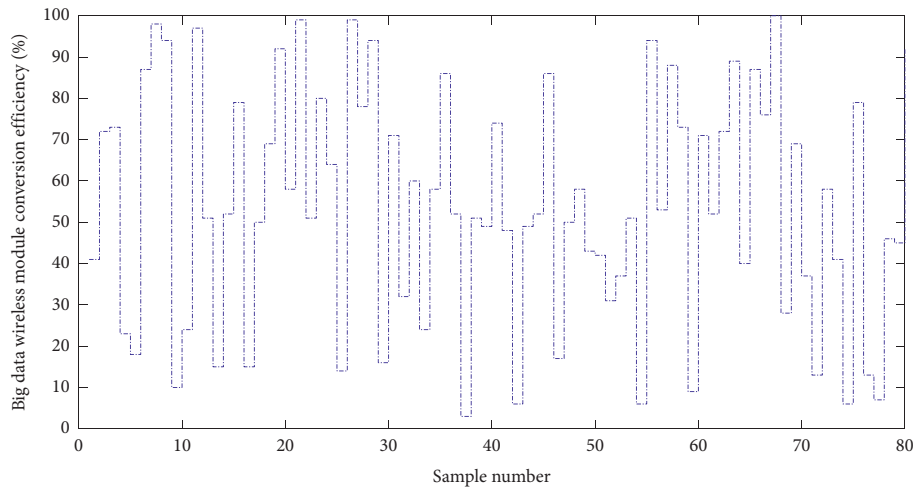
FIGURE 8: Conversion efficiency distribution of big data wireless modules.
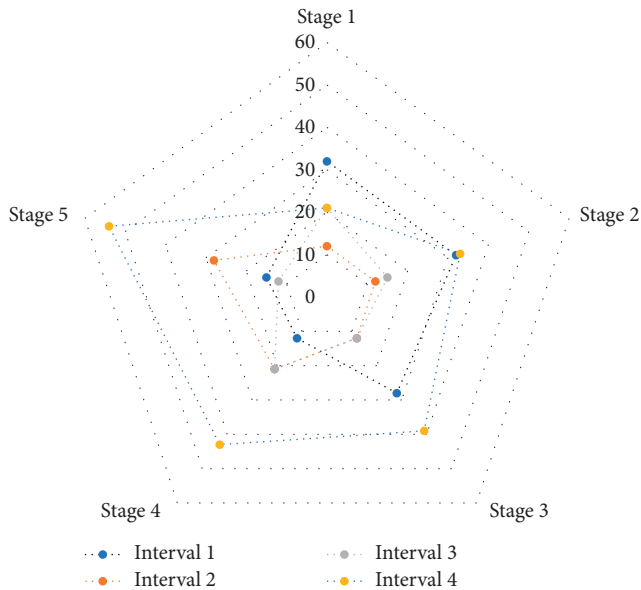
Figure 9: Distribution of news and public opinion serial port data acceptance rate.

propagation speeds to calculate the distance between the two nodes. When querying in the same data table that stores the data records, the system time-consuming will increase with the increase of the query data. When the number of records in the data storage is different but the number of records in the query is the same, the system response time will increase as the total number of storage tables increases.

5. Conclusion

The article analyzes the specific sources of news and public opinion big data, the specific collection methods of data sources, and the methods of various database storage technologies, combined with wireless sensor network technology, open-source big data processing technology, and distributed data storage technology, and it proposes to solve big data. The article mainly works from two aspects of acquisition and storage. In terms of data acquisition, firstly, the application research on the sensor data acquisition network is carried out. Then, the sensor data acquisition system is designed and implemented. Finally, the system is used in news data acquisition and processing. The application was experimented with. In terms of data storage, the article mainly compares the database technologies under big data applications, designs the structure of the HBase data storage system when it is applied to news and public opinion sensor data storage, and conducts an experimental test on its storage performance. The system analyzed the actual needs of the project and the characteristics of big data of news and public opinion Internet of things, combined with the existing big data processing technology, and designed a system that can be used for quick query and analysis of big data of news and public opinion Internet of things. The system is implemented and tested on the Presto framework to achieve the effect of supporting the centralized storage

and rapid query of massive news and public opinion detection data, as well as fit analysis and visual display based on historical data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest in this article.

Acknowledgments

References

[1] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the Internet of Things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[2] H. Fu, G. Manogaran, K. Wu, M. Cao, S. Jiang, and A. Yang, "Intelligent decision-making of online shopping behavior based on Internet of Things," *International Journal of Information Management*, vol. 50, pp. 515–525, 2020.

[3] Z. Alansari, N. B. Anuar, A. Kamsin et al., "Challenges of Internet of Things and big data integration," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 200, pp. 47–55, 2018.

[4] Z. Allam and Z. A. Dhunny, "On big data, artificial intelligence and smart cities," *Cities*, vol. 89, pp. 80–91, 2019.

[5] M. Yazici, S. Basurra, and M. Gaber, "Edge machine learning: enabling smart Internet of Things applications," *Big Data and Cognitive Computing*, vol. 2, no. 3, p. 26, 2018.

[6] A. Mohamed, M. K. Najafabadi, Y. B. Wah, E. A. K. Zaman, and R. Maskat, "The state of the art and taxonomy of big data analytics: view from new big data framework," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 989–1037, 2020.

[7] R. Fernandez Molanes, K. Amarasinghe, J. Rodriguez-Andina, and M. Manic, "Deep learning and reconfigurable platforms in the Internet of Things: challenges and opportunities in algorithms and hardware," *IEEE Industrial Electronics Magazine*, vol. 12, no. 2, pp. 36–49, 2018.

[8] Z. Engin and P. Treleaven, "Algorithmic government: automating public services and supporting civil servants in using data science technologies," *The Computer Journal*, vol. 62, no. 3, pp. 448–460, 2019.

[9] Y. Kabalci, E. Kabalci, S. Padmanaban, J. B. Holm-Nielsen, and F. Blaabjerg, "Internet of Things applications as energy internet in smart grids and smart environments," *Electronics*, vol. 8, no. 9, p. 972, 2019.

[10] J. H. Kim, "A review of cyber-physical system research relevant to the emerging IT trends: industry 4.0, IoT, big data, and cloud computing," *Journal of Industrial Integration and Management*, vol. 2, no. 3, Article ID 1750011, 2017.

[11] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on the cognitive Internet of Things," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 29–35, 2018.

[12] C. Zhang and Y. Chen, "A review of research relevant to the emerging industry trends: industry 4.0, IoT, blockchain, and

business analytics," *Journal of Industrial Integration and Management*, vol. 5, no. 1, pp. 165–180, 2020.

[13] M. M. Losavio, K. P. Chow, A. Koltay, and J. James, "The Internet of Things and the smart city: legal challenges with digital forensics, privacy, and security," *Security and Privacy*, vol. 1, no. 3, p. e23, 2018.

[14] T. Yigitcanlar, N. Kankanamge, and K. Vella, "How are smart city concepts and technologies perceived and utilized? A systematic geo-Twitter analysis of smart cities in Australia," *Journal of Urban Technology*, vol. 28, no. 1-2, pp. 135–154, 2021.

[15] N. Hossein Motlagh, M. Mohammadrezaei, J. Hunt, and B. Zakeri, "Internet of Things (IoT) and the energy sector," *Energies*, vol. 13, no. 2, p. 494, 2020.

[16] J. Chin, V. Callaghan, and S. B. Allouch, "The Internet-of-Things: reflections on the past, present and future from a user-centered and smart environment perspective," *Journal of Ambient Intelligence and Smart Environments*, vol. 11, no. 1, pp. 45–69, 2019.

[17] A. J. A. M. Van Deursen and K. Mossberger, "Any thing for anyone? A new digital divide in Internet-of-Things skills," *Policy & Internet*, vol. 10, no. 2, pp. 122–140, 2018.

[18] M. Jia, A. Komeily, Y. Wang, and R. S. Srinivasan, "Adopting Internet of Things for the development of smart buildings: a review of enabling technologies and applications," *Automation in Construction*, vol. 101, pp. 111–126, 2019.

[19] W. Tushar, N. Wijerathne, W.-T. Li et al., "Internet of Things for green building management: disruptive innovations through low-cost sensor technology and artificial intelligence," *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 100–110, 2018.

[20] L. M. Dang, M. J. Piran, D. Han, K. Min, and H. Moon, "A survey on Internet of Things and cloud computing for healthcare," *Electronics*, vol. 8, no. 7, p. 768, 2019.

[21] F. Behrendt, "Cycling the smart and sustainable city: analyzing EC policy documents on Internet of Things, mobility and transport, and smart cities," *Sustainability*, vol. 11, no. 3, p. 763, 2019.

[22] X. Gong, L. Duan, X. Chen, and J. Zhang, "When social network effect meets congestion effect in wireless networks: data usage equilibrium and optimal pricing," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 449–462, 2017.

Research Article

Comprehensive Evaluation and Optimization Model of Regional Fire Protection Planning of Major Hazard Sources Based on Multiobjective Fuzzy Theory

Fujiang Chen (iD),[1] **Junying Chen,**[2] **and Jingang Liu**[2]

[1]*Xihua University, School of Emergency Management, Chengdu 610039, Sichuan, China*
[2]*China Railway Scientific Research Institute Co., Ltd., Chengdu 610032, Sichuan, China*

Correspondence should be addressed to Fujiang Chen; 1220190024@mail.xhu.edu.cn

The regional fire protection plan is an important part of the city's overall plan, which represents the deepening of the regional fire protection plan and the specific content of the city's economic regional fire protection plan. Fire protection is an important part of national economic and social development, and it is also one of the indicators to measure the level of modern culture. Effective and practical fire protection planning can effectively prevent and reduce fire risks and protect the lives and property of the people, which is very important for social development. In order to optimize the regional fire model, this study uses a very objective fuzzy theory to analyze and discuss the research objects. In view of the large amount of fuzzy information in fuzzy optimization, fuzzy criterion recognition, and fuzzy grouping, based on the generalized fuzzy distance and fuzzy number, this paper proposes a multiobjective fuzzy theory-based comprehensive evaluation and optimization model of fire planning for major hazard sources. The results show that hotel fire risks tend to be higher, which is the focus of people's attention. Among the fire hazards, wholesale, retail, and catering industries are the hardest hit areas, accounting for 66.3%, and some other industries are also disaster areas that need to be dealt with.

1. Introduction

1.1. Background and Significance. With the rapid development of modern social economy and technology, and the great change of energy structure, focus has changed from relying mainly on coal to relying more on the production of petrochemical products. Due to the popularity of the petrochemical industry, many related companies have come here. Petrochemical industry is the cornerstone of national economic development, related to social production and people's life, and plays an important role in promoting social and economic development and improving people's living standards. However, the petrochemical industry is also a very dangerous industry. The accumulation of dangerous and harmful substances has the characteristics of flammability, explosion, high temperature, high pressure, toxic, and harmful. Petrochemical enterprises must pay special attention to fire safety management

in production and operation. If a fire occurs in a certain area, other areas will be affected, resulting in an increase of accident impact and the expansion of the disaster area, which will lead to serious economic loss, social loss, and social consequences. So, it is particularly important to carry out fire planning in high-risk areas (those containing flammable, explosive, high temperature, high pressure, toxic, and harmful items).

1.2. Related Work. Fire station layout planning is an important part of fire protection system. At present, there are many unreasonable and subjective factors in the latheyt planning of domestic fire stations. Wenjuan used the optimization theory design to establish a reasonable framework for the optimization model of the fire station. He transformed the multiobjective fuzzy theory into a unified single-objective model and determined the regional priority and danger level of each part of the

fire according to the fire risk assessment method and transformed it into a genetic algorithm to solve the model [1].

There are always some problems in the mountain fire station. H. Ren takes the planning of the mountain fire station as the research object, mainly aiming to address the shortcomings of the mountain fire planning and the new problems brought about by the rapid development of the city, and identifies and evaluates the fire risk of the buildings in the mountain city. On the basis of fire risk assessment, GIS technology is introduced to solve the problems of location selection of new fire station and latheyt of old fire station. In an area proximal to high fire incidence area, 4 fire stations shall be added. At the same time, the new fire stations are also required to be close to the urban main road as much as possible. The comprehensive application of various methods provides scientific basis and suggestions for the latheyt optimization of municipal fire stations [2].

In the traditional aviation forest protection, the work of cruise and fire location using topographic map and GPS is heavy and difficult, and the time of forest and fire investigation is limited. Jiang analyzed the application of Google Maps in aerial forest protection work. This method is conducive to improving the attention of flight observers in positioning navigation and forest fire investigation and improving the efficiency of aerial forest protection. This paper introduces the functions of Google Earth software in searching destination; acquiring 3D terrain; acquiring landmark; adding annotation; and measuring length, perimeter, and area, and discusses the simulation application of Google Earth, including route planning, firepower positioning, surrounding terrain prediction, etc., according to route navigation, and drawing fire behavior map and building forest fire information database [3].

1.3. Innovation. The purpose of this study was to make a reasonable and effective plan for the regional fire control of major hazard sources, so as to reduce personal injury and economic loss caused by fire accidents. Using the theory of fuzzy set to evaluate the fire risk can provide reasonable early warning for fire protection in major hazard areas. The results of the study of a hotel fire building show that the fire risk of the hotel belongs to a higher risk, accounting for 17% of all fire accidents.

2. Related Concepts

2.1. Multiobjective Fuzzy Theory of the Fire Protection System. In the design of the fire protection system, the models used for fire station latheyt belong to the category of maximum coverage model. The most typical is the problem of MCLP [4]. This model limits the optimal address of gas station to the network node and gives the optimal algorithm in general. The purpose of this model is to find the largest cell group that a server can serve in a normal time or space when the number of servers is limited [5, 6]. However, this model does not consider server congestion and assumes that there are always inactive servers when providing services. In addition, the parameters and structure of these constraint models can be determined [7].

For the fire protection design model, some parameters are not clear and are uncertain, and each part of the whole system has different rules, which means that the objective design problem should be used for analysis and modeling [8]. For example, the time or distance required from a fire station to the scene of a fire is a very important factor in limiting or targeting the function of the MCLP model. However, the traditional MCLP model only deals with time simply, and does not fully consider the difference of time. The degree of fire danger will greatly affect the scientific nature of the site selection decision [9, 10]. This is because different levels of fire need different times to put out the fire. At the same time, the demand points of different risk levels have different fire rescue levels. In the model [11], it is very important to reflect the individual differences between demand points for the optimal design of fire stations. Therefore, conducting a global study requires multiobjective planning [12]. Global research means a thorough study of the fire scene to detect and prevent fires in the first place.

The basic definition of Fuzzy Theory:

If X represents the target set of x, fuzzy set A represents an ordered queue set.

$$A = \left\{ (x, \mu_A \sim (x)) | x \in X \right\}. \tag{1}$$

Here, x represents the element in set X, and $\mu_A \sim (X)$ represents the membership function or membership degree of element X in fuzzy set A, the membership space of mapping x is [0, l].

If

$$\mu_A \sim (\lambda x_1 + (1 - \lambda) x_2) \geq \min \{ \mu_A \sim (x_1), \mu_A \sim (x_2) \}, x_1, x_2 \in X, \lambda \in [0, 1], \tag{2}$$

then fuzzy set A is convex. If $\sup_x \mu_A(x) \approx 1$, then fuzzy set A is called a normal set. The membership function $\mu_c \sim (x)$ of intersection C of fuzzy sets A and B are as follows:

$$\mu_C \sim (x) = \min\{\mu_A \sim (x), \mu_B \sim (x)\}, x \in X,$$
$$\mu_C \sim (x) = \mu_A \sim (x)\mu_B \sim (x), x \in X. \tag{3}$$

So far, the introduction to fuzzy sets is completed.

2.2. Regional Risk Classification.

According to the scope of fire risk and basic characterization, regional risk is divided into regional fixed risk and regional actual risk [13]. Regional fixed risk refers to the impact caused by regional risk and is one of the main sources of accident risk. The actual regional risk is based on the internal risk of the region, taking into account the extension or inhibition of regional factors that affect the risk of accident consequences. The combination of regional risk factors and internal risk indicators is an indicator system to evaluate regional reality [14].

2.2.1. Regional Risk Assessment.

Regional fixed risks are mainly caused by regional accidents, and inherent risks can be quantified by specific methods [15]. At present, the commonly used quantitative analysis methods include fire and explosion index method, six stage risk assessment method, Monde assessment method, inflammable material assessment method, and major hazard source assessment method [16].

2.2.2. Factors Affecting Regional Risk.

Regional risk factors can be divided into risk diffusion factors and risk mitigation factors. Risk diffusion factors refer to the factors that enhance the impact of accident consequences, especially economic vulnerability and social sensitivity [17]. Economic vulnerability refers to the impact on buildings, mechanical devices, residential areas, etc. In post-accident areas, this is usually measured by economic loss indicators [18]. Social sensitivity is often related to psychology, which refers to the psychological damage caused by panic caused by accidents.

Risk mitigation factors can mitigate accident consequences and enterprise emergency management level. The government is an important factor to reduce accident risk [19]. If there is no accident, the assessment of emergency management level is often reflected in the construction of an emergency team and the formulation of an emergency plan. In the case of an accident, the content of emergency management level measurement is mainly the protection of emergency resources available at the scene of the accident and the protection of resources, which is finally reflected in the emergency rescue and evacuation capacity of the area [20].

2.3. Fire Risk Assessment Method.

After long-term development and improvement, the fire risk assessment method has developed to more than 40 kinds, and has been widely used globally. Various quantitative expression forms of evaluation results can be divided into three evaluation methods: quality evaluation method, semiquantitative evaluation method, and quantitative evaluation method [21].

Quality evaluation is the basis for improving product quality. The establishment of a complete quality evaluation system can predict the development trend of the macro quality level; improve the market competitiveness of enterprises; prevent potential dangers and threats to human health, safety, and living environment; and enable people to exchange products, services, and resources with each other. Choose adequately and reasonably in a market economy. The quality assessment method is based on industry standards, regulations, rules, etc., allowing assessment experts with rich theoretical and practical knowledge to draw conclusions based on their observation and experience. The conclusion should consider the following conditions: if there are substances that may cause danger in the area, whether the management within the scope meets the requirements, and the possibility and scope of risk occurrence. The whole evaluation is mainly based on expert reasoning and intuitive judgment, and the mathematical model on which the theory is based is generally simple. This evaluation method has the advantages of simple operation and intuitionistic evaluation conclusion, and has been widely used. However, due to the lack of standards and subjectivity of evaluators in the process of operation, the conclusions drawn are quite different [22, 23].

The semiquantitative evaluation method should identify and distinguish the potential risk factors in the evaluation target according to the risk degree. The basis is: the corresponding classification standard, which is used to classify the determined factors, determines the relevant evaluation indexes, and establishes a link. The link can be a comprehensive mathematical model that defines the relative values of patience and overall safety parameters and quickly assesses the risk level of the object. Semiquantitative method makes up for the shortcomings of randomness and difference of quality method, but it is still simple and intuitive. Fire and explosion indicator method, Monde method, Japan's six stage assessment method, and China's risk classification method are the most commonly used semiquantitative assessment methods.

The quantitative evaluation method includes problem identification, risk identification, cause analysis and impact assessment, risk assessment, and final assessment conclusion. The purpose of the quantitative method is to get more accurate evaluation results. The quantitative evaluation method should be based on the quantification of the evaluation index system. All risk assessment indicators should be consolidated and quantified in accordance with specific criteria. Use a specific mathematical model. Get the risk value of the evaluation target. When analyzing a relatively complex system, the quantitative evaluation method can more clearly reflect the hierarchical relationship between different indicators. The conclusion drawn by this method is very intuitive and clear. We can easily compare the evaluation results of different subjects, or we can compare the evaluation results with the minimum social security standards, which is acceptable to people.

2.4. Develop Effective Plans for Fire Protection in Major Hazardous Areas

2.4.1. The Establishment of Optimization Model of the Fire Water Supply System. In order to ensure that the fire-fighting system provides effective fire-fighting measures and reduces the losses caused by the fire, the design of the fire water supply system is a vital link in the fire protection planning. Whether the regional fire-fighting water supply system is designed to be reliable, reasonable, and economical will affect the scientific nature of the entire regional fire-fighting plan [24]. It can be seen from the previous norms and standards that the existing fire-fighting water consumption norms are mainly based on the population size of the city to determine the frequency of fires, thereby determining the fire-fighting water consumption of the city. However, for the design of regional fire protection planning, this kind of fire water consumption index based solely on population size is extremely unreasonable. For example, for an industrial area that contains some major hazards like liquefied petroleum gas tanks, the population may be relatively small in the area, but once a fire occurs in this area, the amount of fire extinguishing water required will be greater than that of a pure living area with the same population [25, 26]. Therefore, if the population of an area is solely used as the basis for determining the water volume of the fire water supply system, the design standard will be far lower than the water required for the actual fire, which will not effectively achieve the purpose of disaster reduction. Therefore, in order to comprehensively analyze the influencing factors of regional fire water consumption and use the interrelationship of various factors to plan the fire water consumption of the entire regional water supply system, it is necessary to adopt suitable optimization methods to establish a reasonable mathematical model of the fire water supply system. Specific influencing factors include water consumption, distance from the fire station, water storage, floor height, and flow of people. Its purpose is to realize the scientific rationality of the planning of the fire water supply system to meet the needs of actual production and life to achieve the purpose of disaster reduction and elimination.

2.4.2. The Latheyt of Each Functional Block Is Reasonable. It should be noted that the fire hazard in the LPG operation area is high and there are many accidents. Once a fire occurs, it is extremely destructive. Although the fire accident in the LPG operation area may not directly affect the national strategic oil reserve base, it directly affects the petrochemical operation area, has an indirect impact on the national strategic oil reserve base, and increases the overall fire risk. From the perspective of the overall situation and long-term interests, it is possible to consider limiting the scale of LPG operation areas.

2.4.3. Fire Station Latheyt. One of the main responsibilities of the fire department is to fight fire and rescue. This task usually comes from the arranged fire station. The fire station is the basic organization unit of the fire department for fire-fighting and rescue. The purpose of planning a fire station is to enable the fire station to provide effective fire-fighting and rescue services. Based on the analysis of the current situation of the fire protection environment, the number, location, area of responsibility, floor space, building area, vehicles, equipment, and personnel to be planned for the construction of fire stations are clearly defined, and the time to achieve these goals are determined. There are specific plans for human and material resources [27]. The planning of the fire station must adapt to the needs of multifunctional and large-area operations. For areas with special fire protection requirements, a fire station capable of handling special fires and disasters should be planned and established. If possible, the three-dimensional latheyt of water, land, and air should be considered for comprehensive development [28].

3. Experimental Design

3.1. Research Object. As a hotel has a large flow of people, a lot of flammable materials, and high floors, improper management can easily lead to huge fires. Once a fire occurs, it is a huge accident. So in this study, a hotel on fire is selected as the object, and the fuzzy set theory is used to evaluate the fire risk. The evaluation index system includes influencing factors, key variables, and hierarchy system. The dynamic and fuzzy characteristics of the system are reflected in the evaluation of fire risk, fire resistance, and fire vulnerability. In addition, this method is also a kind of evaluation method, which is widely used, practical, and verified by a large number of practices.

3.2. Research Steps. The fuzzy evaluation method has been widely used in fuzzy system evaluation for its good practicality. It uses an ambiguous theory and examines the use of multiple indicators to assess participation. Three key groups participated in the whole process, including a set of systematic factors, a set of systematic reviews, and a set of evaluation factors. Among them, the evaluation factors have the greatest influence and ultimately affect the evaluation results of credibility. According to the fire risk assessment index system, the system factor set $U = \{U_1, U_2, U_3\}$, wherein U1 represents the fire risk degree, U2 represents the fire vulnerability degree, and U3 represents the fire resistance degree. Determine the weight of system coefficient; each factor includes a set of evaluation indexes, and the total weight A of this factor is determined according to the weight of evaluation index a1, wherein $A = \{a_1, a_2, a_3\}$, $\sum_{i=1}^{n} a_1 = 1$.

According to the fire protection level, it is divided into five evaluation levels. The evaluation set $V = \{V_1, V_2, V_3, V_4, V_5\}$, V1 is very safe, V2 is relatively safe, V3 is general safety, V4 is relatively dangerous, and V5 is very dangerous. The specific evaluation process is as follows:

Through 10 experts to grade the second-level indicators under the first-level indicators, grade and summarize the second-level indicators under the fire risk U1 of a hotel. The fuzzy subsets of fire risk can be obtained, and then U2 fire

vulnerability and U3 are the fuzzy subsets A2 and A3 of fire resistance:

(1) These experts need to score the first-level indicators of fire risk U1, fire vulnerability U2, and fire resistance U3, and synthesize the results,

The fuzzy subset $A = \{A_1, A_2, A_3\}$ of fire risk factors is calculated by using the formula.

(2) Factor evaluation. 30 experts were invited to grade the indicators in the indicator system according to the specific conditions of the experimental subjects and the criteria of indicator quantification. For example, for an indicator, 13 experts think it is very safe, 11 experts think it is safe, and 6 experts think it is safe. Then, the evaluation set of this factor is defined as {0, 47, 0, 42, 0, 11, 0, 0}. The sum of all evaluation sets is 1. This rule also applies to other indicators, and then summarizes the evaluation value of each individual factor.

$R_1 = A_1 * r_1, R_2 = A_2 * r_2, R_3 = A_3 * r_3$, get the fuzzy matrix $R = \{R_1, R_2, R_3\}$, through the formula $B = A * R$, according to the scoring of experts, finally determine the fire risk level.

4. Regional Fire Protection Planning of Major Hazard Sources

4.1. Fire Risk Assessment. The fire risk assessment process is closely related to the fire development process. In general, the specific development process can be divided into the following processes: fire, growth, full development, recession, and finally extinction. In the above experiment, a hotel is taken as the object, and the fuzzy set theory is used for fire risk assessment and analysis. At the same time, the level of fire risk assessment in recent years can be found by querying the National Statistical Yearbook. The results are shown in Table 1.

According to the data provided in Table 1, we can clearly see that relative safety accounts for half of the fire risk rating of the experts. In fire vulnerability, most experts choose the fire is very safe. In the fire resistance industry, most experts give the scores of very safe and relatively safe. It can be seen that the fire risk is the most likely factor to increase the fire risk, followed by fire vulnerability and fire resistance.

4.2. Response Time of Fire Accidents. Using many objective and uncertain methods to solve the problem of optimizing the latheyt of fire stations has a simple feature: adaptability. Combined with the analysis and research on the level of fire risk assessment and the reference statistics of some countries, the fuzzy characteristics can be better determined. As shown in Table 2, it is the recommended fire response reference value of the fire protection system.

The rescue time of fire-fighting is calculated in seconds. The earlier the rescue team arrives at the scene of the accident, the faster the control of the accident will be. Table 2 also shows that different risk levels have different response times. The average operating speed can be used to turn the requirement of operating time into a distance limit. Risk levels should be used to distinguish multiple limits on the distance from the fire station to the scene of the accident. The calculation of this risk level should be based on the statistics of different regions. Fire risk assessment and analysis can be used to obtain data required for different risk levels. The number of fire vehicles and firefighters shall be determined according to the latest regional statistical accident report and fire station budget.

4.3. Data Analysis of Major Hazard Accidents. Relevant data show that in 2007, five of the top ten fires outside China occurred due to major flammable and explosive hazards in urban public places and facilities. Although the causes and consequences of accidents are quite different, major accidents involving hazardous chemical sources have some common characteristics: we ought to prevent accidents, which not only leave a large number of casualties inside and outside the factory but also cause huge property losses, and eventually damage the environment with immediate effect. At present, China has not established a systematic and effective major risk supervision system; the latheyt, category, and dynamic status of major hazard sources are unclear; and the ability of supervision, control, early warning system, and emergency rescue to prevent emergencies has not been established, so major accidents happen repeatedly.

As shown in Figure 1, major safety accidents occur frequently. The main reason is that the location of major hazard sources in most cities is not planned, the classification is not clear, the effective supervision system of major risk sources cannot be established, the emergency rescue plan cannot be determined, and the demand and ability of emergency dispersion cannot be determined. In particular, the goal of macro-control management is not clear. Therefore, how to avoid the occurrence of major accidents and how to implement regular protective measures are the key issues in realizing emergency rescue. It is necessary to strengthen the awareness of fire control and implement fire control measures, and ensure timely and effective treatment for major accident victims.

4.4. Characteristics of Major Hazard Industries. The research field of major hazard sources is defined within the scope of cities.

Typical units of major risk sources are production sites (such as chemical companies and heating facilities companies) and storage sites (such as oil, explosives, and

TABLE 1: Fire risk assessment.

Fire risk	Very safe	Relatively safe	General safety	More dangerous	Very dangerous
Fire hazard	7	15	8	0	0
Fire vulnerability	12	8	7	3	0
Fire resistance	16	12	2	0	0

TABLE 2: Fire accident response time.

Risk level	Time limit to reach the accident location	Distance from the accident site (km)	Response time to fire accidents (s)
I	3–6	3–6	35.5
II	6–10	6–10	33.2
III	11–13	11–13	29.6
IV	14–20	14–20	15.4
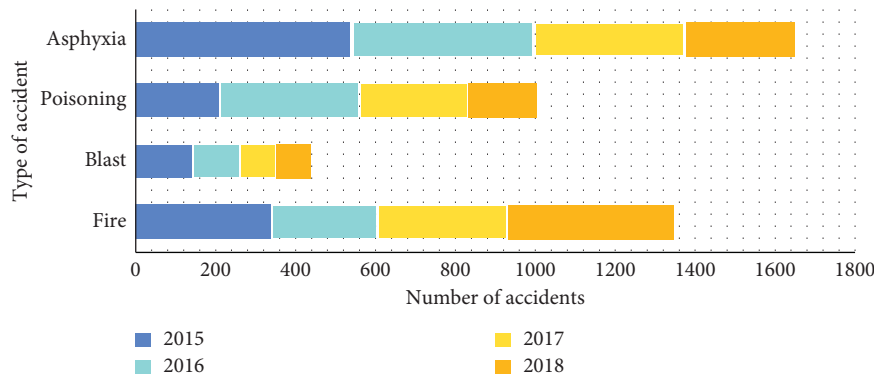
FIGURE 1: Hazard area accident.

toxic substances storage), which are distributed in areas in different urban buildings (such as gas filling stations and LPG station). Because the direct correspondence between important risk sources and land types is not easy to obtain, the main risk sources of land use types can be obtained only by defining the characteristics of important risk sources in the industry and the relationship between major risk industries and land use types (Figure 2).

4.5. Fire Water Consumption. In order to effectively implement fire rescue, it is necessary to evaluate the initial water consumption of each fire hazard source. Only by using the strong theoretical basis determined by the estimation of fire water consumption, can we truly plan a reasonable fire water supply system.

As shown in Figure 3, the calculation results of water consumption will be different for different estimation methods. Among them, the water consumption calculated by the New Zealand method is the most effective than that calculated by the other two methods, namely, the Iowa method and ISO method.

4.6. Regional Fire Protection Planning. As a typical regional disaster causing factor, the main risk sources have different effects on regional risk. The distribution characteristics of the main risk sources are not equal to the spatial distribution, and in the absence of geographic information, they do not even belong to the category of spatial distribution. This study considers that major hazard sources have important characteristics in industrial distribution and urban land-use distribution. Regional risks are to be analyzed from this perspective. Especially when analyzing the main risk distribution characteristics of urban land, in addition to considering its own land, the distribution characteristics of different types of land affected by major risks should be fully analyzed according to different urban models. All urban land use information can be considered as major disaster source.

As shown in Figure 4, in the absence of regional details, the main distribution of major risk sources should be adapted to the main land-use types and their characteristics. Combined with the characteristics of urban latheyt, especially the characteristics of land latheyt, preliminarily
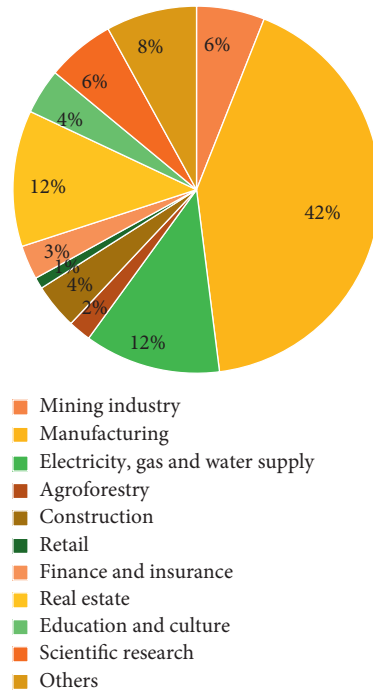
FIGURE 2: Characteristics of major hazard industries.
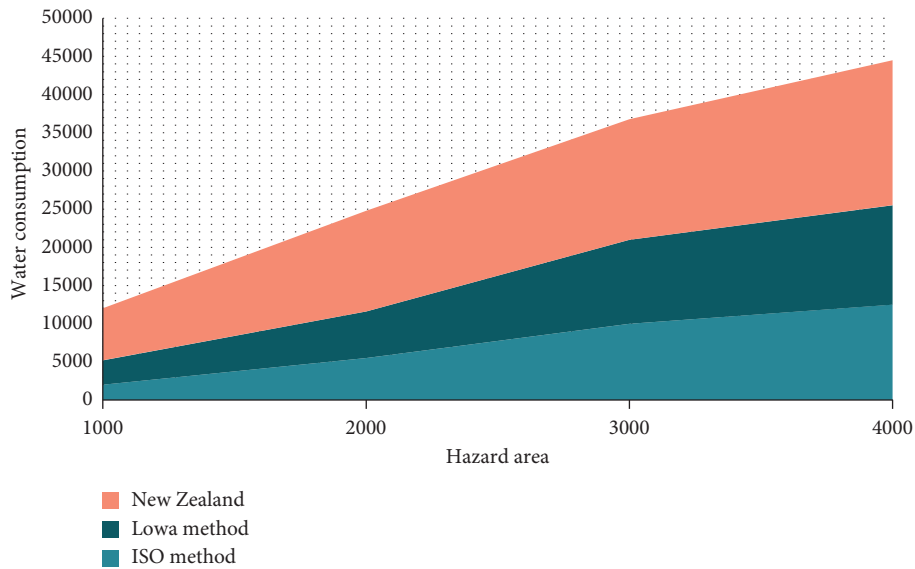
FIGURE 3: Fire water consumption comparison.

analyze the impact of major risks on different regions, divide the site into different types of land, and analyze the land types that may be affected by major risks. After the sites on various types of land are adjusted, the vulnerability analysis of these people's activity sites can be used as a perspective to analyze the risk distribution in the region. Setting up fire-fighting facilities in places with high risk can greatly reduce the destructive power of fire.

FIGURE 4: Regional fire protection planning.

5. Conclusions

Multiobjective fuzzy decision-making is still a booming research field, and there are many theoretical and practical problems to be solved. The main work of this study focuses on the multiobjective fuzzy theory and its related fields, and some research results have been achieved. At the same time, there are many projects that need to be further studied and improved. The qualitative, objective, and weight of fuzzy set theory is one of the main research contents, which is based on the consistency of a series of projects and objectives. In most cases, this makes sense. However, there may be two other situations in the decision-making process. First, the priority given by the decision-maker does not conform to the principle of consistency; second, the decision-maker cannot give sufficient ranking. In other words, the sequencing of some projects or goals is incomplete or unknown. The fuzzy optimization model is applied to group decision-making, which further widens the application field of fuzzy sets, but at the same time, the aggregation of group fuzzy preferences also faces the problem of ordering consistency, that is, in addition to the inconsistency of individual preferences, there is also the inconsistency of group preferences. It is not only an interesting and realistic problem but also a common problem in group decision-making.

This paper is based on the multiobjective fuzzy theory, a comprehensive evaluation and optimization model for the fire protection planning of major dangerous source areas. The establishment of this model firstly analyzes the various influencing parameters of the fire station layout. The model not only considers conventional response time constraints but also economic constraints such as fire protection investment, disaster mitigation losses, and even all possible future situations. The waiting time of the fire station service and the major hazards that need to be considered make the layout optimization strategy of the regional fire station more suitable for the development needs of the society in the future. Based on national standards and successful cases abroad, this paper adopts multiobjective fuzzy planning theory and establishes a regional fire water supply system planning model based on major hazards. In the model, the

area is first divided into several functional blocks. Calculate and analyze the fire-fighting water consumption of each functional block, and then consider the special requirements of major hazard sources of fire-fighting water consumption and the premise that the water supply system of each functional area can cooperate and assist, and then determine the fire-fighting water from the overall consideration of the area, in order to plan out the fire water supply system of the area.

According to the data of fire risk, fire resistance, and fire sensitivity, this study determines the state of disaster, analyzes the characteristics of major hazard sources, carries out risk assessment and emergency device investigation, analyzes the impact of chemical hazard sources on the site, and evaluates and optimizes the spatial latheyt of emergency sites. Detailed inspection of traffic conditions, land use, land characteristics, fire-prone areas, and other factors shall determine the candidate locations for fire planning latheyt. From the two levels of the government and the individual, the fairness goal, the economic cost goal, and the distance cost goal of the fire planning latheyt are set up, and the corresponding decision-making model is constructed.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] W. Zhang, H. Ma, J. Zhang, L. Chen, and Y. Qu, "Multiobjective fuzzy-optimization of crowbar resistances for the low-voltage ride-through of doubly fed induction wind turbine generation systems," *Journal of Power Electronics*, vol. 15, no. 4, pp. 1119–1130, 2015.

[2] H. Ren, Y. Lu, Q. Wu, X. Yang, and A. Zhou, "Multi-objective optimization of a hybrid distributed energy system using NSGA-II algorithm," *Frontiers in Energy*, vol. 12, no. 4, pp. 518–528, 2018.

[3] W. Jiang and J. Min, "Design of winding parameters based on multiobjective decision-making and fuzzy optimization theory," *Journal of Control Science and Engineering*, vol. 2019, no. 11, pp. 1–8, 2019.

[4] B. Wang, Y. Li, S. Wang, and J. Watada, "A multi-objective portfolio selection model with fuzzy value-at-risk ratio," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 6, pp. 3673–3687, 2018.

[5] B. Gao and J. Wang, "Multi-objective fuzzy clustering for synthetic aperture radar imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2341–2345, 2015.

[6] A. Trivedi and A. Singh, "A hybrid multi-objective decision model for emergency shelter location-relocation projects using fuzzy analytic hierarchy process and goal programming approach," *International Journal of Project Management*, vol. 35, no. 5, pp. 827–840, 2017.

[7] M. Singh and J. S. Dhillon, "Multiobjective thermal power dispatch using opposition-based greedy heuristic search," *International Journal of Electrical Power & Energy Systems*, vol. 82, no. nov, pp. 339–353, 2016.

[8] H. B. Tolabi, M. R. Shakarami, R. Hosseini, and S. B. M. Ayob, "Novel FGbSA: fuzzy-Galaxy-based search algorithm for multi-objective reconfiguration of distribution systems," *Russian Electrical Engineering*, vol. 87, no. 10, pp. 588–595, 2016.

[9] F. Perez and T. Gomez, "Multiobjective project portfolio selection with fuzzy constraints," *Annals of Operations Research*, vol. 245, no. 1-2, pp. 7–29, 2016.

[10] H. Mavalizadeh, A. Ahmadi, and A. Heidari, "Probabilistic multi-objective generation and transmission expansion planning problem using normal boundary intersection," *IET Generation, Transmission & Distribution*, vol. 9, no. 6, pp. 560–570, 2015.

[11] D. Grechi and A. Biggeri, "Polycyclic aromatic hidrocarbons deposition in the Milazzo-Valle del Mela (Sicily Region, Southern Italy) high-risk area following an oil refinery fire," *Epidemiologia e Prevenzione*, vol. 40, no. 1, p. 16, 2016.

[12] S. Modugno, H. Balzter, B. Cole, and P. Borrelli, "Mapping regional patterns of large forest fires in Wildland-Urban Interface areas in Europe," *Journal of Environmental Management*, vol. 172, no. 1, pp. 112–126, 2016.

[13] R. Bhagawat, Z. Lifu, K. Hamidreza, S. Xueijian, and R. Sushila, "Quantifying the spatiotemporal pattern of urban expansion and hazard and risk area identification in the kaski," *District of Nepal Land*, vol. 7, no. 1, p. 37, 2018.

[14] W. Schwanghart, J. Seegers, and G. Zeilinger, "Landslide susceptibility mapping in three selected target zones in Afghanistan," *Journal of Biomechanics*, vol. 39, no. 06, p. S42, 2015.

[15] L. I. Yuanyuan, Z. Zhuqun, L. I. Si, L. Yujia, and L. Wenxing, "Research on the electromagnetic environment of ships based on the high frequency algorithm Chinese," *Journal of Ship Research*, vol. 12, no. 2, pp. 202–208, 2015.

[16] Y. Wang, R. Wang, L. Fan et al., "Assessment of multiple exposure to chemical elements and health risks among residents near Huodehong lead-zinc mining area in Yunnan, Southwest China," *Chemosphere*, vol. 174, no. may, pp. 613–627, 2017.

[17] F. Gutiérrez, I. Fabregat, C. Roqué et al., "Sinkholes in hypogene versus epigene karst systems, illustrated with the hypogene gypsum karst of the Sant Miquel de Campmajor Valley, NE Spain," *Geomorphology*, vol. 328, no. MAR.1, pp. 57–78, 2019.

[18] P. E. Zope, T. I. Eldho, and V. Jothiprakash, "Hydrological impacts of land use-land cover change and detention basins on urban flood hazard: a case study of Poisar River basin, Mumbai, India," *Natural Hazards*, vol. 87, no. 3, pp. 1267–1283, 2017.

[19] L. Nahayo, C. Mupenzi, G. Habiyaremye et al., "Landslides hazard mapping in Rwanda using bivariate statistical index method," *Environmental Engineering Science*, vol. 36, no. 8, pp. 892–902, 2019.

[20] A. Dietrich and M. Krautblatter, "Evidence for enhanced debris-flow activity in the northern calcareous alps since the 1980s (plansee, Austria)," *Geomorphology*, vol. 287, no. jun.15, pp. 144–158, 2017.

[21] L. Milosevic, E. Mihajlovic, and G. Janackovic, "Novel approach to landfill fire protection engineering based ON multi-criteria analysis and principles OF sustainable," *ENVIRONMENTAL MANAGEMENT Journal of Environmental Protection & Ecology*, vol. 19, no. 1, pp. 226–235, 2018.

[22] W. Meyn, "Untersuchungen zum Bestandsschutz von Holztreppen aus brandschutztechnischer Sicht am Berliner Baurecht," *Bautechnik*, vol. 94, no. 11, pp. 812–820, 2017.

[23] L. Yin, W. Pan, J. Kuang, and M. Zhuang, "Application of bootstrap-DEA with fuzzy computing in performance evaluation of forklift leasing supplier," *IEEE Access*, vol. 8, pp. 66095–66104, 2019.

[24] A. Takakusagi and M. Sudo, "Study on failures and troubles about security systems, fire protection systems, manlifts and container vehicles in hospital facility," *AIJ Journal of Technology and Design*, vol. 23, no. 54, pp. 597–602, 2017.

[25] M. Klippel and A. Frangi, "Brandverhalten von Brettsperrholz," *Bautechnik*, vol. 93, no. 8, pp. 567–573, 2016.

[26] G. Khatwani and P. R. Srivastava, "Impact of information technology on information search channel selection for consumers," *Journal of Organizational and End User Computing*, vol. 30, no. 3, pp. 63–80, 2018.

[27] T. Coleman, P. Sorensen, T. F. Yaeger, and G. Hinrichs, "A strategic planning intervention employing large group change: a scholar/practitioner application," *Academy of Management Proceedings*, vol. 2016, no. 1, Article ID 10278, 2016.

[28] M. Hu, Yi Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers in Neuroscience*, p. 965, 2021.

Research Article

An Intelligent Music Production Technology Based on Generation Confrontation Mechanism

Yanjing Li [1] **and Xinyuan Liu** [2]

¹Department of Music, Langfang Normal University, Langfang, Hebei 065000, China
²Department of Electronic and Information Engineering, Langfang Normal University, Langfang, Hebei 065000, China

Correspondence should be addressed to Yanjing Li; 1321464@lfnu.edu.cn

In recent years, with the development of deep neural network becoming more and more mature, especially after the proposal of generative confrontation mechanism, academia has made many achievements in the research of image, video and text generation. Therefore, scholars began to use similar attempts in the research of music generation. Therefore, based on the existing theoretical technology and research work, this paper studies music production, and then proposes an intelligent music production technology based on generation confrontation mechanism to enrich the research in the field of computer music generation. This paper takes the music generation method based on generation countermeasure mechanism as the research topic, and mainly studies the following: after studying the existing music generation model based on generation countermeasure network, a time structure model for maintaining music coherence is proposed. In music generation, avoid manual input and ensure the interdependence between tracks. At the same time, this paper studies and implements the generation method of discrete music events based on multi track, including multi track correlation model and discrete processing. The lakh MIDI data set is studied. On this basis, the lakh MIDI is pre-processed to obtain the LMD piano roll data set, which is used in the music generation experiment of MCT-GAN. When studying the multi track music generation based on generation countermeasure network, this paper studies and analyzes three models, and puts forward the multi track music generation method based on CT-GAN, which mainly improves the existing music generation model based on GAN. Finally, the generation results of MCT-GAN are compared with those of Muse-GAN, so as to reflect the improvement effect of MCT-GAN. Select 20 auditees to listen to the generated music and real music and distinguish them. Finally, analyze them according to the evaluation results. After evaluation, it is concluded that the research effect of multi track music generation based on CT-GAN is improved.

1. Introduction

As an important way of expression in the field of art, music embodies a series of human unique thinking modes and is a unified combination of regularity and creativity [1]. On the one hand, the composition of music is naturally based on certain music theory rules, such as melody, rhythm, mode, chord, harmony, polyphony, musical form [2]. Music that cannot meet the constraints of music theory rules is often poor in auditory sweetness and cannot be accepted by the public [3]. On the other hand, music that simply meets the constraints of music theory is not necessarily good music. Music itself also acts as an important task of emotional

expression carrier [4]. This requires the Creator not to stack the rules in a conventional way, but to incorporate innovation into the music, so that the generated melody will not be stereotyped and the mode is fixed. At the same time, the automatic generation of music by computer algorithm has always been a field of concern. Music creation with the help of computer algorithm can reduce the production threshold and save manpower and time cost [5]. To a certain extent, avoid the copyright problem, and quickly make a large number of music according to the scene needs, such as a large number of customized melodies that render the plot emotion in film and television dramas [6]. Due to the regularity and creativity of music and other artistic creation

fields, the research on intelligent music creation can well measure and test the performance of artificial intelligence. When describing the experimental results of their own model algorithms, many related works in the field of music generation use the test method of organizing volunteers for auditory recognition [7]. The investigation and statistics are carried out from the aspects of authenticity, ear pleasing, creativity and interest. For example, the literature randomly looks for 255 testers to test the quality of music samples generated by their Muse-GAN model [8]. Similarly, some other algorithmic composition work also adopts this human judgment method. Generally speaking, the production process of intelligent music needs to minimize the workload of human intervention [9]. This production method can generate music automatically or semi automatically. The output results should not only meet the basic prior knowledge of music theory, but also have some algorithm creativity. The literature discusses that this process is to independently produce continuous audio signals or discrete symbol sequences from the computational model, and these signals and sequences must meet the music theory architecture [10]. As early as the 1950s, artificial intelligence technology was just in its infancy. Although limited by data and hardware performance, people also began to explore the field of intelligent composition and achieved some results [11]. In the early stage, intelligent music was mainly generated in two ways. First, it was created based on statistical analysis and combined with Markov chain models. For example, the literature was the first to create String Quartet suite with large-scale computer, which became the first music work completely generated by computer in history [12]. It used Markov chain model to generate random notes with limited control, combined with the rules of harmony and polyphony, these notes are tested, and the tested elements are modified to synthesize the string quartet of traditional music notation. The second is simple pattern matching and machine learning based on music theory rules, such as integrating music theory into machine learning to produce notes [13]. In recent years, deep learning has attracted more and more attention. Researchers have frequently applied deep learning technology in many fields and achieved good results. Therefore, some fields have produced projects that can be put into practical use. Based on this, people's research on neural network is emerging again. With the improvement of computer hardware technology, the deep learning technology based on neural network can not only deal with huge data samples, but also have strong computing power. Because of this, a series of problems such as many model parameters and difficult training have been solved, which are widely used to solve problems in various research fields. For example, deep learning technology greatly improves the accuracy of image classification, even exceeds human classification ability, and has successful applications in the field of natural language processing. At present, the generative countermeasure network based on neural network has been widely used in the fields of image, vision, voice and language. Therefore, this paper proposes to realize the automatic annotation of music based on generative countermeasure network.

With the continuous evolution and updating of AI technology, the formal technology of music creation has been further developed [14]. In recent years, generative countermeasure network has attracted more and more attention in the field of data generation. Relying on its strong fitting ability and simple training derivation process, it has been introduced into many application fields [15]. Generative countermeasure network has been widely used in graphics generation, graphics compression, speech generation and super-resolution restoration. The literature points out that "GAN can also generate text, dialogue generation, machine translation, voice generation, etc. At the same time, GAN is also involved in other fields, such as music generation, password decoding [16]. However, the application effect of GAN in other fields is not significant, so how to improve the application effect of GAN in other fields will be worthy of in-depth research, so as to make the generated countermeasure network shine in artificial intelligence " [17]. The research on speech synthesis and music creation using GAN has also been in the stage of exploration and development, and many achievements have been made. Through the investigation, it is found that the number of literature review on algorithmic composition is relatively large, but the literature review on music generation based on deep learning, especially based on generating confrontation network, is relatively missing.

2. Related Work

As early as the 1990s, the literature first proposed the idea based on confrontation generation. The author trained the predictor to judge the input data mode, and let each data minimize its predictability, forming the simplest confrontation competition learning mode. In 2014, P. Shamsolmoali [18] et al. formally put forward the concept of generating countermeasure network, using the confrontation competition of generator model and discriminator model to realize semi supervised learning, which opened up a new field for the research of data generation. The biggest difference between GANs and the traditional generation model is that in the process of data training, it is both unified and antagonistic. The optimization directions of generators and discriminators are different from each other, forming a competitive relationship, but their optimization calculation depends on each other's output to form a unified system [19]. In the confrontation training mode, the generator no longer directly learns the distribution from the training data set, but indirectly iteratively learns through the optimization direction given by the discriminator to generate fake samples that confuse the false with the true. Compared with the traditional unsupervised learning models such as self coding and autoregression, GANs has the advantages of fast calculation speed, better sample quality, strong expansion flexibility and so on [20]. Generally speaking, the production process of intelligent music needs to minimize the workload of human intervention. This production method can generate music automatically or semi automatically. The output results should not only meet the basic prior knowledge of music theory, but also have some algorithm creativity. The

literature discusses that this process is to independently produce continuous audio signals or discrete symbol sequences from the computational model, and these signals and sequences must meet the music theory architecture.

Subsequently, a large number of theoretical and technical research results of generative countermeasure networks came out one after another, and some of them played a milestone role in promoting the overall research progress of generative countermeasure networks. The verification model of the original GANs is realized by multi-layer perceptron MLP, and the generation quality is poor [1]. The literature proposes DCGANs, and the generator and discriminator are realized by deep convolution network respectively, so as to ensure the engineering implementation of GANs in the field of graphics generation. In order to improve the training stability of GANs and make the output data have a certain controllable directivity, the literature proposes the conditional generation confrontation model CGANs [21]. On this basis, the literature adds a supervised learning classification task to GANs to form ACGANs to improve the generation quality of the model. In order to further explore the training stability of GANs, some targeted training skills are added in the training process of GANs to form improved GANs. The literature improves the loss function of GANs from the mathematical principle, so that GANs can better narrow the distribution of generated data and training set data [22]. This research work has played an important role in the development of GANs technology, and also made a series of application achievements in many task fields, especially in the direction of graphics and images. At present, there are many research works on GANs with wide coverage. This paper investigates and classifies the existing achievements from the key technical level and application level. It should be noted that the same work may contain improvements in multiple directions [23].

There are some problems caused by the loss function in the training process of the original GANs. For example, it is proved by theory that the original loss function will cause the gradient to disappear when the coincidence between the generated data distribution and the training set data distribution is not enough [24]. This problem cannot be improved by simple model structure optimization. Therefore, some work has improved the loss function, so as to improve the generation quality of GANs. Firstly, it is proved that the distance measurement between distributions based on KL distance and JS distance is discontinuous, which leads to the instability of discriminator training. The sample data generated by GANs is constrained by confrontation training and tries to fit the distribution of real data sets [25]. However, there is no guidance of conditional control in the whole training process, and the randomness of generated data is great, which also indirectly leads to the instability of training. The literature integrates conditional control into countermeasure training, uses conditional vectors to control the attributes of generated samples, and indirectly introduces the decoupling input of conditional variables. On the basis of CGANs, the literature uses discriminators for supervised learning, classifies samples, further enhances conditional constraints, and further improves the quality of generated samples [26]. The proposal of generative countermeasure network not only presents a new network model architecture, but also provides a framework of generative countermeasure training idea. After a certain expansion of the framework, it can be combined with other existing model methods to generate samples for specific fields.

3. Generative Confrontation Mechanism and Its Important Derivative Model in the field of Music Generation

3.1. Generation Confrontation Mechanism. The generative countermeasure mechanism consists of two sub networks: generator and discriminator. The task of generator F is to make the model fit the real training data distribution as much as possible, and discriminator B is committed to distinguishing whether the input data comes from real data or false data manufactured by the generator. In the training process, they continuously improve their ability, and finally achieve an ideal state of Nash equilibrium, so that the data generated by the generator is as close to the real data distribution as possible. The basic loss function of GAN is:

$$\min_{B} \max_{F} V(B, F) = \frac{T_{Z-Pz(z)}[\ln(1 - B(F(z)))]}{E_{x-(z)} \int [\prod D(x)] \mathrm{d}x}. \tag{1}$$

In theory, the generator and discriminator can be realized by any differentiable function, which is not limited to multi-layer perceptron or convolutional neural network. GAN and its derived music generation model are shown in Figure 1.

4. Generation of Derivative Models of Countermeasure Networks

4.1. WGAN. Zhang [21] et al. have studied and discussed various problems existing in the original generated countermeasure network. There are two main reasons for these problems in the original generated countermeasure network: 1. Unreasonable distance measurement of equivalent optimization (KL divergence and JS divergence); 2. There is no intersection or overlap between the data distribution generated after the random initialization of the generator and the actual data distribution, which can be ignored. However, it is not suitable to use JS divergence to measure the distance between the disjoint parts. Theoretically, this is also the main reason for the instability of the original generation anti network training. Therefore, Wasserstein GANs was put forward by homeopathy (WGAN). WGAN has made a good improvement on the problem of unstable training of generation countermeasure network model. WGAN uses Wasserstein distance as an alternative to JS divergence to measure the distance between generated data distribution and real data distribution, and solves the two problems of unstable training and unable to provide numerical indicators to measure the training process. Compared with KL divergence and JS divergence, Wasserstein distance can still be used to measure the distance between the generated data distribution and the real data distribution when there is no intersection or the intersection part can be ignored.
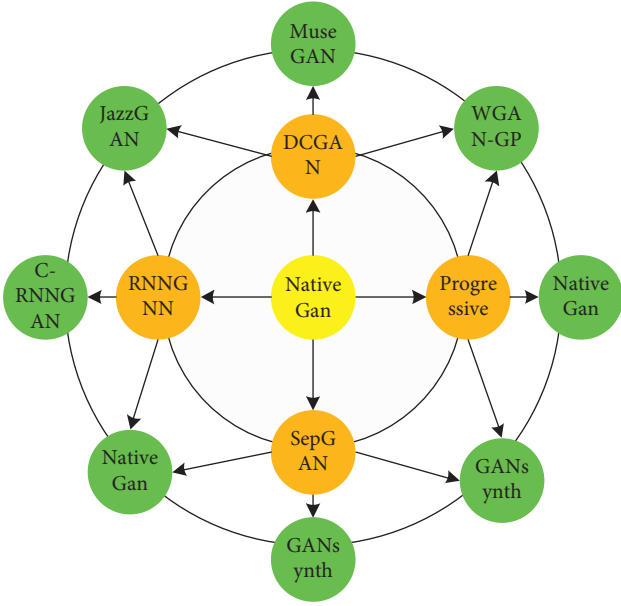
FIGURE 1: Generation model of GAN and its derived music.

Before introducing the basic principle of WGAN, we first need to understand the concept of Lipschitz continuity, that is, for a continuous function $f(x)$, if there is a constant $k > 0$, any two elements x and $x0$ in the definition domain meet:

$$\int_{x_0}^{x} g(x) + g(x_0) \leq F\sqrt{(x - x_0)^2 - 4x} + \sqrt{x^2 - 4x}. \quad (2)$$

Wasserstein distance can not only reflect the distance of two nonoverlapping or negligible overlapping distributions, but also provide a meaningful gradient, that is, it is smooth. However, it is impossible to solve the lower bound directly. Therefore, the Wasserstein distance can be changed into the following form according to Kantorovich Rubinstein duality:

$$\min_{B} \max_{F} \prod \left[E \int (x) \right] + \frac{\sum\limits_{x-p} E(\widetilde{X})}{\left| x - \widetilde{X} \right|}. \quad (3)$$

4.2. WGAN-GP. In view of the problems of unstable training and slow convergence in the generated countermeasure network, Wasserstein GANs proposed by Yang [22] et al. has made a good improvement, but sometimes the generated samples are still bad or the convergence fails. Then, scholars such as Shi [23] proposed that the main reason for the failure of bad sample generation or convergence based on WGAN is that the weight clipping method is used to meet the Lipschitz continuity condition in the implementation of WGAN, which will lead to bad training behavior. In this regard, scholars such as Ishaan Gulrajan proposed WGAN-GP (improved training of Wasserstein GANs), that is, the improvement of WGAN network model based on gradient penalty, which uses gradient penalty instead of weight clipping to meet Lipschitz continuity condition. In this case, the extreme distribution of parameter weights cannot give full play to the generalization ability of deep neural network. In addition, weight shearing can easily lead to gradient explosion or gradient dispersion. The reason is that when weight shearing is carried out, if the shear threshold is set too small, the gradient will gradually decrease through each layer of network, and then decay exponentially until the gradient disappears; When the shear threshold is set too large, the gradient will gradually increase after each layer, and then the gradient will explode, as shown in Figure 2.

Therefore, only when the threshold is set at a balanced position can the normal gradient return be guaranteed. However, due to the narrow balance area, the parameter adjustment work becomes very troublesome.

4.3. CT-GAN. Aiming at the problems of difficult training and slow convergence of WGAN, WGAN-GP improves WGAN and uses the gradient penalty method instead of the weight clipping method to meet the Lipschitz continuity condition. WGAN-GP makes network training faster and convergence faster, but WGAN-GP also has defects. According to what is mentioned in the paper, the difference between gradient penalty and weight clipping is that gradient penalty can never be punished to every place through gradient penalty term in a limited number of training times, and the result is that gradient penalty term GP only works on sampling point x. The important parts of the support set cannot be checked at all, especially the observed data points and the basic manifold supporting the real data distribution pr.

In addition, CT-GAN does not only focus on one specific data point at a time, but sets regularization on a pair of data points drawn near the manifold according to the most basic definition of Lipschitz continuity. In particular, CT-GAN perturbs each actual data point x twice, and uses the Lipschitz constant to limit the difference between the response of the discriminator to the set data point.

Figure 3 describes the main idea of CT-GAN. Because the gradient penalty often fails to verify the continuity of the region near the real number x, the discriminant function can freely violate the Lipschitz continuity. Therefore, on the basis of gradient penalty, CT-GAN uses the inspection of two disturbed Lipschitz continuity conditions near any observed actual data point, so as to alleviate the problem of WGAN-GP.

4.4. Generative Model and Discriminant Model. The main structure of the generator model (discriminator model) is shown in Figure 4. It can be seen that its structure is relatively simple. The main reason is the repeated iterations in the training process. Each round of training is equivalent to training dozens of generators and discriminators, and each iterative operation needs to store the neural network parameters of the intermediate process, so only one layer of neural network is used here. Moreover, there is a problem of parameter feedback in the training process of generator and discriminator. The trained discriminator parameters are superimposed into the generator parameters as negative feedback. Too complex network will lead to poor feedback process.
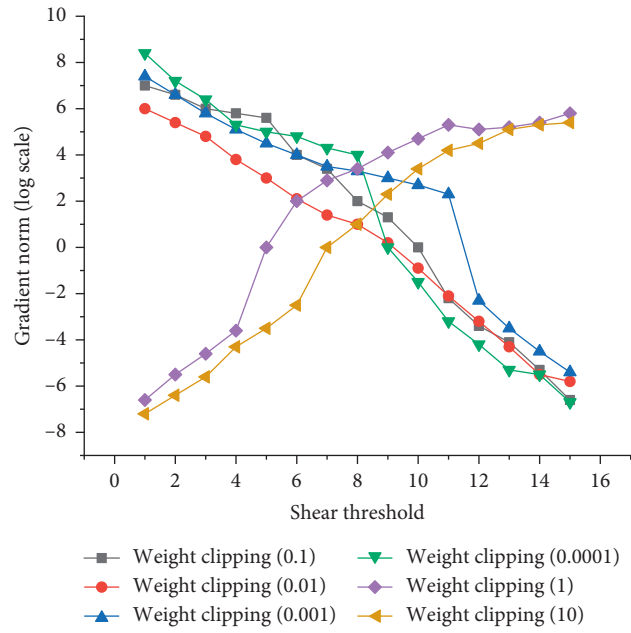
FIGURE 2: Shear threshold gradient experimental diagram.
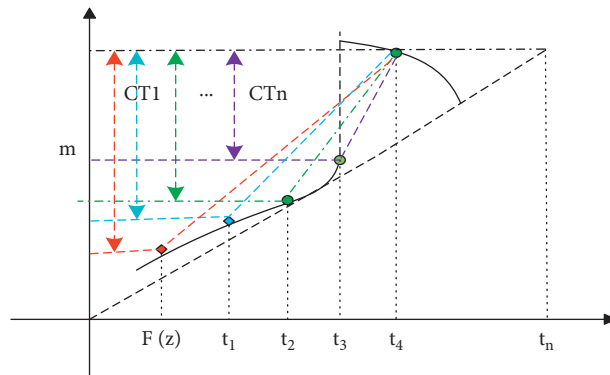
FIGURE 3: Main idea diagram of CT-GAN.
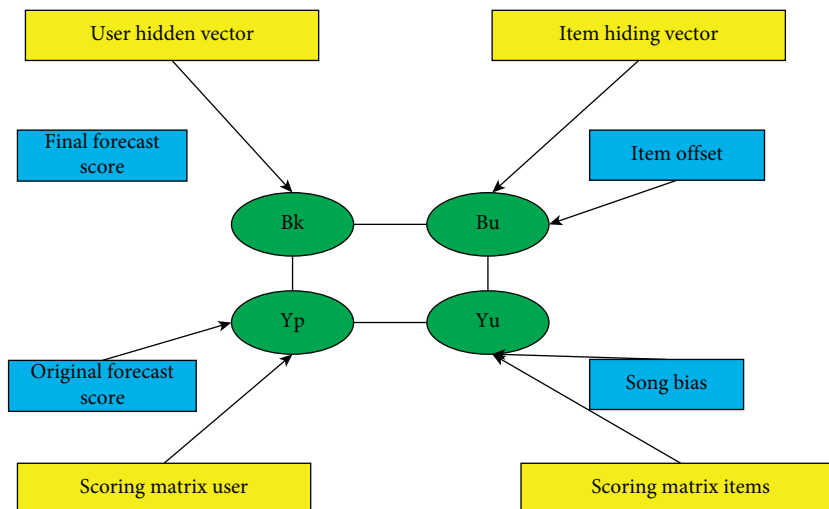
FIGURE 4: Generator simulation.

5. Generation Experiment and Analysis of Intelligent Music Generating Confrontation Mechanism

5.1. Data Set and Pre-Processing. In the experiment of music generation in this paper, the MIDI file of lakh MIDI data set (LMD) is selected to generate music containing the following five tracks: bass, drum, guitar, piano and strings.

LMD is used because it provides a rich and real collection of MIDI files and some related metadata. The lakh MIDI dataset is a collection of 176581 unique MIDI files, of which 45129 have been matched and aligned with the entries in the Millennium Song Dataset. The goal of Lakh MIDI dataset is to facilitate large-scale music information retrieval, including symbols (only MIDI files) and audio content-based (using information extracted from MIDI files as annotations for matching audio files). Lakh MIDI data sets have many different versions, some of which are subsets and some of which are obtained after processing. LMD-full is the complete set of Lakh MIDI data sets. Each file is named according to its MD5 checksum. However, since invalid MIDI files are not deleted from this collection, it contains thousands of files that may have been damaged; LMD matched is a subset of LMD-full containing 45129 files, which matches the entries in the Millennium Song Dataset.

However, in this experiment, the music in piano roll format is generated instead of MIDI format, so Lakh MIDI data set will not be directly used as the training data set of this multi track music generation experiment, but a series of pre-processing steps will be performed on Lakh MIDI data set before training, Thus, the LMD piano roll data set needed by the experiment is obtained, and then it is put into the network for training to generate multi track music. In the next two sections, we will first introduce the data representation of this experiment-Piano Roll format, and then introduce the pre-processing process of data set.

5.2. Data Pre-Processing. According to experience, there are some defects in using LMD directly for music generation, mainly because of the mapping between track and MIDI channel. Because these MIDI files are usually user generated, the mapping is error prone. Therefore, before the experiment, the following three data cleaning steps are performed to pre-process the data.

Divide the different tracks of each MIDI music file in LMD: This paper divides each MIDI file into five tracks: Bass, drum, guitar, Piano and string. The music generated by the experiment also includes these five tracks. This is because there is no clear method to identify and distinguish the track where the main melody and chord are played. Therefore, the music track cannot be divided into melody, rhythm and drum.

Select "rock" songs and merge tracks with sparse notes. Firstly, this paper mainly selects the music marked as "rock" songs, because this category generally contains the above five tracks, which can reduce the number of blank sections in these tracks. Although songs with genre tags are filtered, some tracks still tend to play only a few notes in the whole

song. This increases data sparsity and hinders the learning process. This paper combines the tracks with sparse notes into the most similar one of the above tracks. Because string instruments are the most diverse among a few instruments, instruments that are difficult to classify are merged into the track where string instruments are located.

5.3. Experimental Results and Analysis. After training, a large number of music clips are generated using the MCT-GAN model in this paper. After listening, it is found that most of the generated music is smooth and has a certain artistic aesthetic effect. Here, randomly select the generated music clips as the result display. Firstly, the generated multi track music is presented in the format of score. The generated music file is converted from npz format to mid file format by using Python library pypianoll, and then the mid file is expanded in the format of music score by using Cubase software. In order to make the music score clearly displayed, only four sections of music fragments are presented, and separate music scores are derived for each section. Therefore, the music score includes four sections of generated music.

As mentioned above, MCT-GAN has made two innovations in the previous music generation model. 1. Use CT-GAN as the generation confrontation model. 2. Propose a new time structure model. Muse GAN is a multitrack music generation network based on WGAN-GP. Therefore, this paper compares the MCT-GAN experiment with Muse-GAN experiment, so as to analyze the experimental effect of MCT-GAN proposed in this paper. According to the experimental results, the generation results obtained by MCT-GAN and Muse-GAN experiments are compared as follows:

Firstly, the training loss of the discriminator at different training iteration points is obtained experimentally. Therefore, the convergence curve of discriminator loss can be obtained and compared with the convergence curve of Muse GAN. The results are shown in Figure 5.

From the two convergence curves of Muse GAN and MCT-GAN in Figure 5, it can be found that compared with Muse-GAN based on WGAN-GP, MCT-GAN with increased CT penalty term converges faster and the training process is more stable.

Through the experiment, the distribution histograms of discriminator parameters in the two experimental iterations can be obtained respectively, as shown in Figure 6.

From the comparison in the figure, it can be found that the interval range of the parameter distribution obtained from the music generation experiment based on MCT-GAN is significantly smaller than that obtained from the Muse-GAN experiment, which shows that compared with the Muse-GAN based on WGAN-GP. The MCT-GAN with consistency penalty proposed in this paper is less prone to over fitting. In conclusion, through experimental comparison, MCT-GAN is better than Muse-GAN in stability, convergence speed and over fitting. These indicators can be calculated from the training (actual) data and the generated data to obtain the corresponding indicator values, and by comparing the values calculated from the actual data and the generated data, the effect of

(a)

(b)
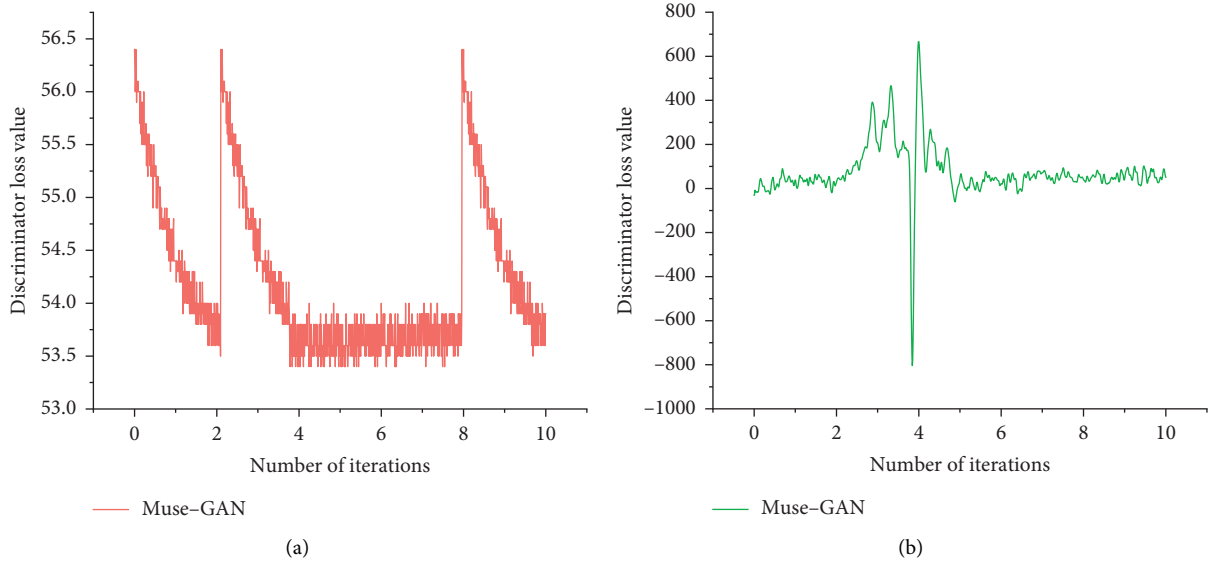
FIGURE 5: Comparison diagram of convergence curve of discriminator loss.
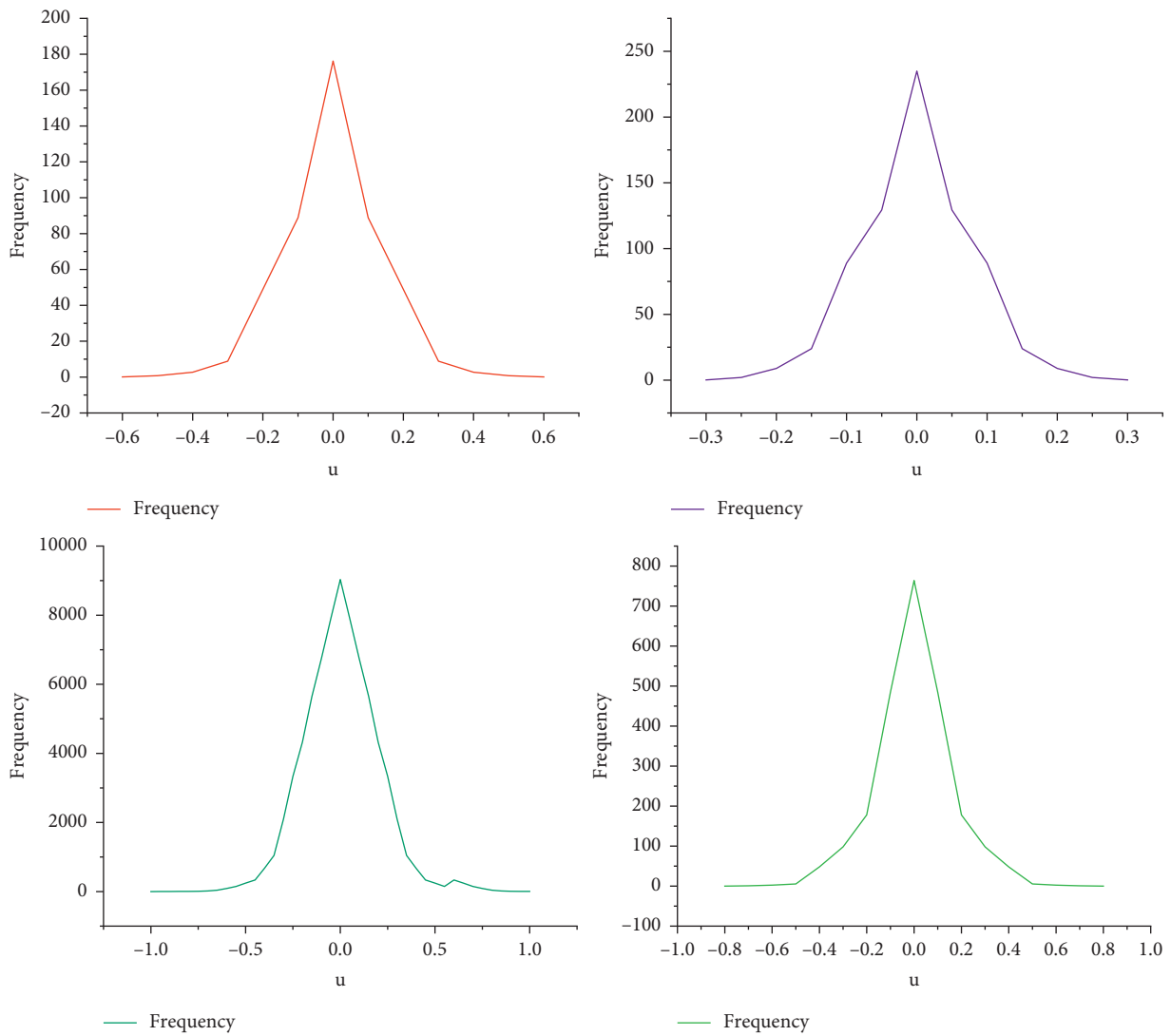
FIGURE 6: Parameter distribution histogram.
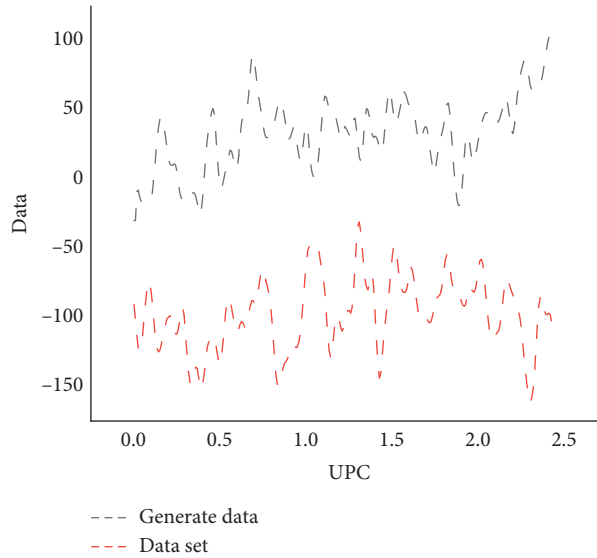
Figure 7: UPC index curve of stringed instruments.

(a)

(b)
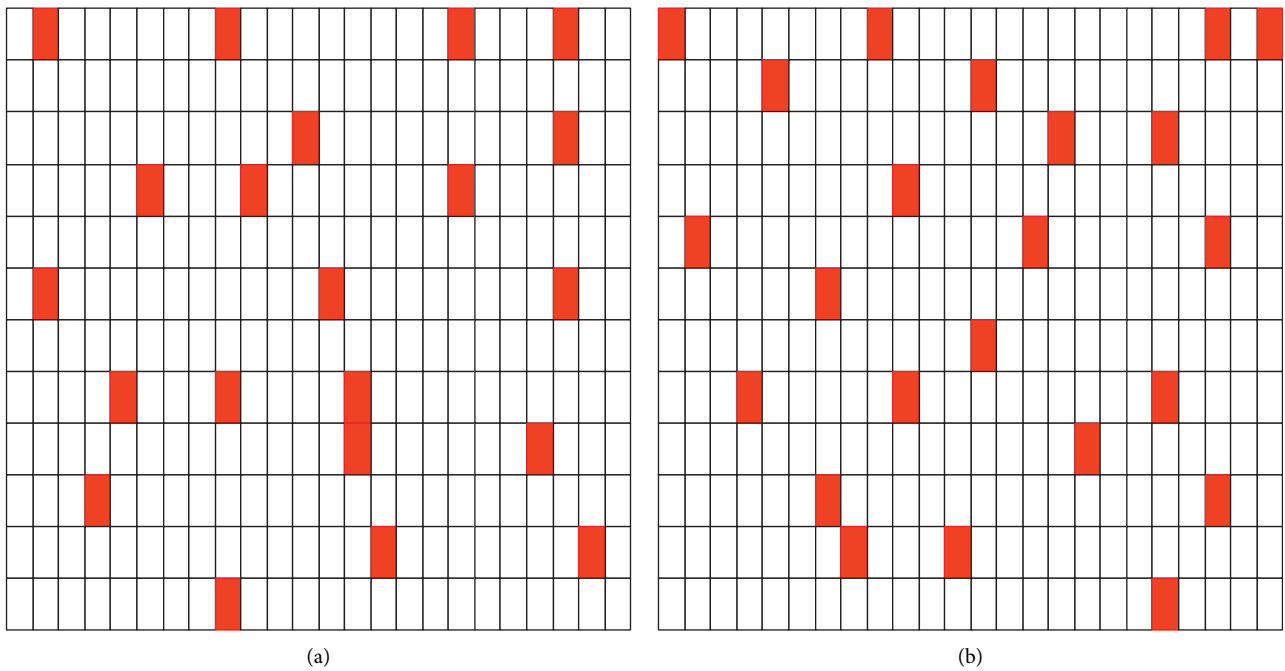
Figure 8: Comparison of blank sections.

the generator can be understood. Therefore, the idea of the index evaluation method is consistent with the idea that the distribution of real and fake data in GAN should be close.

Muse-GAN designed an index evaluation method to evaluate the generated results. Therefore, this paper uses this method to compare Muse-GAN and MCT-GAN. The following indicators are defined in Muse-GAN, where EB, UPC and QN are in track evaluation indicators, and TD is inter track evaluation indicators. These indicators can be calculated from the training (actual) data and generated data, and the effect of the generator can be understood by comparing

the values calculated from the actual data and generated data. Therefore, the idea of index evaluation method is consistent with the idea that the distribution of real and false data in GAN should be close.

Figure 7 shows the change of UPC index value of the track where the string instrument is located during the experiment. The blue line in the figure is the UPC index value of the string instrument in the training set. It can be found from the figure that G can learn this index during training, which also shows that these index values can be used to preliminarily evaluate the generation results of the model before subjective evaluation.

In addition, this paper also randomly selects several sections of piano volumes as a comparative display, as shown in Figure 8. Figure 8a is a piano volume diagram for generating music using T3 as the time structure, and Figure 8b is a piano volume diagram for generating music using T1 as the time structure. The sections framed by the red box (or thick box) in the figure represent empty sections. From the comparison of the two figures, it can be seen that there are fewer blank sections using T3 as the time structure than using T1, which shows that the generation effect of MCT-GAN model using T3 as the time structure is improved compared with Muse-GAN.

In conclusion, using the index evaluation method to compare the experimental results, it is found that compared with the music generation experiment based on Muse-GAN. The MCT-GAN proposed in this paper has improved the effect of generating blank bars, generating note quality and generating music harmony. In addition to comparing the results, this paper also uses a subjective manual evaluation method. Select 80 test listeners and listen to 8 music clips for each tester. The specific proportion of the two is random. Then, after listening to each piece of music, the tester asked them to identify the source of the music and generate the music from the given three options (A. music; *b*. Real music; *c*. Finally, according to the selection results of all testers, the accuracy rate (i.e. the probability that the selection item is correct) and the confusion rate (i.e. the probability that *c* is selected) of each audience are obtained.

The results showed that 72.5% of the testers could not distinguish the generated music from the real data, that is, the number of people with a correct rate of less than 50%. Moreover, among the testers whose accuracy is in the range of 0%–25%, most of the testers cannot distinguish after listening to the music, and finally choose the uncertainty C; After the test, when interviewing the testers whose accuracy is in the range of 75%–100%, it is found that some testers have guessing behavior in discrimination. Therefore, from the results of manual evaluation, it can be found that the music production method proposed in this paper is effective.

6. Conclusion

With the increasing improvement of material living standards, people's cultural and entertainment life is becoming richer and richer, and people's demand for spirit is increasing day by day. However, in the field of artificial intelligence, generating real and aesthetic works of art has always been considered a challenge. With the rapid development of artificial intelligence technology and the development of deep learning technology in recent years, great progress has been made in image, audio and text generation. In particular, the generation experiment is carried out with the very popular generation countermeasure network, and remarkable results are achieved. Therefore, this paper studies the existing theoretical results. Finally, based on the improved derivative model CT-GAN of generation countermeasure network, the generation experiment of intelligent music is completed.

Because the experimental research design of this paper is based on the popular deep learning model to generate countermeasure network, this paper first briefly introduces the development background and basic concepts of deep learning. Then, the basic principle of generating countermeasure network is understood in detail, and the advantages and disadvantages of generating countermeasure network are analyzed. Then, understand an improved model Wasserstein GANs proposed for the problems of generating countermeasure network, and learn its basic principle. Although WGAN has improved Gan network, there are still some problems. Therefore, this paper also studies WGAN-GP, which is an improved model of WGAN network based on gradient penalty. For better training effect, this paper also makes a detailed study on the latest improved model CT-GAN based on consistency penalty term, which is proposed to solve the problems of WGAN-GP.

Learning and improving the existing music generation model based on GaN. When studying the multitrack music generation based on generation countermeasure network, this paper studies and analyzes three models. Based on them, the multitrack music generation method based on CT-GAN is proposed in this paper, which mainly improves the existing music generation model based on GAN. The music generated in this paper only contains five tracks: bass, drum, guitar, piano and string. However, some music creations in real life contain more different tracks. Therefore, in the follow-up work, we can consider using richer music data sets to generate music containing more tracks, so as to make the generated music more diverse. This paper realizes the multitrack music generation experiment based on CT-GAN, which can generate 4 bar music including 5 tracks. In the follow-up work, this music generation can be extended to music composition related applications to assist composers or music lovers in more convenient and diverse music creation. During data pre-processing, in order to reduce the number of empty sections contained in music data and deal with the data imbalance problem that increases data sparsity and hinders the learning process due to several notes played in some tracks in the whole song, the corresponding piano volumes are added and similar musical instruments are combined into one of the above tracks. However, this will bring noise to the data and affect the generation effect.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there is no conflict of interest.

Acknowledgments

References

[1] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.

[2] W. Tang, S. Tan, B. Li, and J. Huang, "Automatic steganographic distortion learning using a generative adversarial network," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1547–1551, 2017.

[3] J. Ma, H. Xu, J. Jiang, X. Mei, and X.-P. Zhang, "DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion," *IEEE Transactions on Image Processing*, vol. 29, pp. 4980–4995, 2020.

[4] S. Xie, Z. Yu, and Z. Lv, "Multi-disease Prediction Based on Deep Learning: A Survey," *Computer Modeling in Engineering & Sciences*, vol. 127, no. 3, pp. 1–34, 2021.

[5] F. Zhu, F. Ye, Y. Fu, Q Liu, and B Shen, "Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network," *Scientific Reports*, vol. 9, no. 1, Article ID 6734, 2019.

[6] T. M. Quan, T. Nguyen-Duc, and W. K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.

[7] V. Volz, J. Schrum, and J. Liu, "Evolving mario levels in the latent space of a deep convolutional generative adversarial network," *Proceedings of the genetic and evolutionary computation conference*, vol. 23, pp. 221–228, 2018.

[8] R. Qian, R. T. Ta, W. Yang, J. Su, and J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, vol. 11, pp. 2482–2491, Salt Lake City, UT, USA, June 2018.

[9] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, "Stock market prediction based on generative adversarial network," *Procedia Computer Science*, vol. 147, pp. 400–406, 2019.

[10] T. Iqbal and H. Ali, "Generative adversarial network for medical images (MI-GAN)," *Journal of Medical Systems*, vol. 42, no. 11, Article ID 231, 2018.

[11] H. Hukkelås, R. Mester, and F. Lindseth, "Deepprivacy: a generative adversarial network for face anonymization," *International Symposium on Visual Computing*, Springer, vol. 13, pp. 565–578, , Heidelberg, Germany, 2019.

[12] J. Song, T. He, and L. Gao, "Unified binary generative adversarial network for image retrieval and compression," *International Journal of Computer Vision*, vol. 128, no. 8, pp. 2243–2264.

[13] R. Sipio, M. F. Giannelli, and S. K. Haghighat, "DijetGAN: a generative-adversarial network approach for the simulation of QCD dijet events at the LHC," *Journal of High Energy Physics*, vol. 110, no. 8, pp. 1–17, 2019.

[14] T. Li, R. Qian, and C. Dong, "Beautygan: instance-level facial makeup transfer with deep generative adversarial network," *Proceedings of the 26th ACM international conference on Multimedia*, vol. 15, pp. 645–653, 2018.

[15] O. Prykhodko, V. S. Johansson, and P. C. Kotsias, "A de novo molecular generation method using latent vector based generative adversarial network," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 1–13, 2019.

[16] X. Zhou, W. Liang, K. I. K. Wang, R. Huang, and Q. Jin, "Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 1, pp. 246–257, 2021.

[17] G. Zhang, M. Kan, and S. Shan, "Generative adversarial network with spatial attention for face attribute editing," in *Proceedings of the European conference on computer vision (ECCV)*, vol. 21, pp. 417–432, Munich, Germany, September 2018.

[18] P. Shamsolmoali, M. Zareapoor, R. Wang, D. K. Jain, and J. Yang, "G-GANISR: gradual generative adversarial network for image super resolution," *Neurocomputing*, vol. 366, pp. 140–153, 2019.

[19] T. Vu, M. Li, H. Humayun, Y. Zhou, and J. Yao, "A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer," *Experimental Biology and Medicine*, vol. 245, no. 7, pp. 597–605, 2020.

[20] A. Sharma and G. Hamarneh, "Missing MRI pulse sequence synthesis using multi-modal generative adversarial network," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1170–1183, 2019.

[21] H. Zhang, Z. Le, Z. Shao, H. Xu, and J. Ma, "MFF-GAN: an unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion," *Information Fusion*, vol. 66, pp. 40–53, 2021.

[22] B. Yang, X. Li, Y. Hou et al., "Non-invasive (non-contact) measurements of human thermal physiology signals and thermal comfort/discomfort poses -A review," *Energy and Buildings*, vol. 224, Article ID 110261, 2020.

[23] Y. Shi, K. Davaslioglu, and Y. E. Sagduyu, "Generative adversarial network for wireless signal spoofing," *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*, vol. 12, pp. 55–60, 2019.

[24] M. N. Kamel Boulos, Z. Lu, P. Guerrero, C. Jennett, and S. Anthony, "From urban planning and emergency training to Pokémon Go: applications of virtual reality GIS (VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health," *International Journal of Health Geographics*, vol. 16, no. 7, pp. 1–11, 2017.

[25] M. Jahanbakht, W. Xiang, L. Hanzo, and M. R. Azghadi, "Internet of underwater things and big marine data analytics – a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 904–956, 2021.

[26] Z. Cai and Z. He, "Trading private range counting over big IoT data," in *Proceedings of the 39th IEEE International Conference on Distributed Computing Systems (ICDCS)*, Dallas, TX, USA, July 2019.

Hindawi

Review Article

An Adaptive Learning Image Denoising Algorithm Based on Eigenvalue Extraction and the GAN Model

Feng Wang ⓘ**, Zhiming Xu, Weichuan Ni** ⓘ**, Jinhuang Chen, and Zhihong Pan** ⓘ

Guangzhou Xinhua University, Guangzhou, China

Correspondence should be addressed to Zhihong Pan; zhpan@xhsysu.edu.cn

This paper proposes a self-adjusting generative confrontation network image denoising algorithm. The algorithm combines noise reduction and the adaptive learning GAN model. First, the algorithm uses image features to preprocess the image and extract the effective information of the image. Then, the edge signal is classified according to the threshold value to suppress the problem of "excessive strangulation," and then the edge signal of the image is extracted to enhance the effective signal in the high-frequency signal. Finally, the algorithm uses an adaptive learning GAN model to further train the image. Each iteration of the generator network is composed of three stages. And then, we get the best value. Through experiments, it can be seen from the data that the article algorithm is compared with the traditional algorithm and the literature algorithm. Under the same conditions, the algorithm can ensure the operating efficiency while having better fidelity, and it can still denoise at the same time. The edge signal of the image is preserved and has a better visual effect.

1. Introduction

The generation of image information will inevitably be more or less accompanied by the generation of noise. Noise is a factor that hinders human sense organs from understanding the received source information. We can define it as an unpredictable, random error. It will have a certain impact on the collection, input, and processing of image information and the final output results. Especially in the process of input, collection, and transmission of image information, if the input is accompanied by large noise, it will definitely have an adverse effect on the subsequent processing process and the processing result. The purpose of denoising is to improve the image and solve the interference problem of the actual image.

With the continuous development of technology, human access to information is obtained through human vision, hearing and touch, and other senses, the vast majority of which information is derived from human vision. And in real life, image acquisition is vulnerable to external interference to form noisy images, and in the process of image segmentation and parameter estimation of noisy images, it will cause errors in the resulting image so that the

image denoising processing will become the current research hot spot in the field of image processing [1–5]. In 1992, Donoho proposed the small wave threshold atrophy method; the algorithm with its own denoise superiority quickly attracted people's attention, but its tendency to "overstrangling" wavelet coefficient and cannot optimally represent the line and surface singularity in the image, so that the wavelet transformation in the image denoising has certain limitations [6, 7], but denoising in the past, there is still the problem of fidelity. With the development of artificial intelligence technology, multimodal fusion is an inevitable trend, and multimodal fusion can make more effective use of the characteristics of the image. In practice, a large amount of labeled data is difficult to obtain, but GAN can solve this problem very well [8, 9]. Therefore, the problem of how to combine GAN and noise reduction came into being.

This paper proposes an image denoising algorithm, which is based on a model that generates an adaptive adjustment against the network. It combines noise reduction and the adaptive learning GAN model, preprocesses the image by combining image characteristics, increases the

effective information quantity of the image while reducing the amount of image calculation, and then uses the GAN model with adaptive learning to further denoise the image, so as to obtain the best noise removal effect. In order to avoid the phenomenon of "overstrangling" in the process of noise reduction, the referenced threshold classifies its edge signal and then extracts the edge signal of the image to enhance the effective signal in the high-frequency signal while deleting the negligible signal in the image. Each iterative update of the generator network λ consists of three stages. In the end, we obtain the key areas in the image and self-learn to obtain the best value for the individual.

2. Generative Adversarial Networks

The generative adversarial network (GAN) was proposed by Goodfellow, which uses convolutional neural networks to train image samples [10, 11]. As a probability generation model, the generation against the network has been applied to many visual tasks, especially in the excellent performance of the image generation direction. The network structure is generated as shown in Figure 1.

The purpose is to let the generator generate data similar to the real data distribution to achieve a fake effect. In terms of data prediction, the training generator introduces images into the generation of the neural network and, through training, reduces the difference between the two so that the distributor cannot distinguish the authenticity of the generated data. Since the difference between the generated data distribution P_G and the real data distribution P_{Data} cannot be calculated, therefore, the discriminator uses a data sampling method to count the divergence of the two and then distinguish them.

The objective function is

$$V(G, D) = E_{x \sim P_{\text{data}}}[\log D(x)] + E_{x \sim P_G}[\log(1 - D(x))]. \quad (1)$$

Among them, $E(*)$ is the mathematical expectation value, and $D(*)$ is the result of the differential output. The objective function of maximizing judgment is the dispersion of generated data and real data. According to the distribution dispersion, the prediction function can be obtained:

$$G^* = \arg\min_G \max_D V(G, D). \quad (2)$$

GAN can generate generated samples similar to the real data distribution, but there are difficulties in training, data cannot be obtained, difficult to converge, and pattern collapse, and some studies have proposed variations of the GAN to improve the model. When there is no overlap between the generated data and the real data distribution, the dispersion of the two data distributions is constant, Log2, which can lead to the disappearance of gradients and make it difficult for the model to train.

Broadly speaking, generators and judges are not enemies, and they are not confrontational relationships. In fact, they are constantly updating and learning on the basis of each other to achieve each other. During the training process, the goal of generating network G is to generate as many real pictures as possible to deceive network D. D is opposite to G, and the goal of D is to distinguish between false images generated by G and real images as much as possible. The two constitute a dynamic "game process."

The success of the GAN also benefits from the success of confrontational ideas. The idea of confrontation has been introduced into several fields including machine learning and artificial intelligence. The two behaviors of "game" and "competition" contain the characteristics of confrontation. People combine game theory and machine learning to produce game machine learning.

The advantages and disadvantages of the GAN are as follows: compared with other generative models, GAN has the following four advantages:

(a) From the actual results, GAN seems to produce better samples than other models.

(b) Most other frameworks require the generator network to have a certain functional form, for example, the output layer is Gaussian. It is important that all other frameworks require that the generator network be distributed with nonzero quality. The generative confrontation network framework can train any type of generator network and can learn to generate points only on thin manifolds close to the data.

(c) Any generator network and any discriminator will be useful because they do not need to design a model that follows any type of decomposition.

(d) There is no need for reasoning during the learning process and no need to use the Markov chain for repeated sampling (inference), thus avoiding difficult approximate calculation problems of probabilities.

Compared with VAE, deep Boltzmann machine, NICE, and real NVE, GAN has no lower limit of change. At the same time, compared with deep Boltzmann machines, GAN does not have tricky partition functions. GAN can be generated all at once through samples, instead of repeatedly applying Markov chain operators.

3. The Algorithm of This Article

At present, fuzzy method training using synthetic data cannot effectively distinguish real images. In order to solve this problem, this paper proposes a new method of noise reduction and adaptive learning GAN model, that is, by combining image characteristics to preprocess the image so that the target is clear, then use the GAN model with adaptive learning to further denoise the image, and then get the best noise removal effect.

At present, many researchers have proposed many denoising methods for images. However, some denoising methods will blur the details of the image while denoising the image, making the original clear image blurred and unable to be restored. Some are able to get clear images, but the amount of calculation is too large to meet people's needs. Nowadays, the most common wavelet denoising algorithms are the modulus maximum denoising algorithm, correlation denoising algorithm, and threshold denoising algorithm.
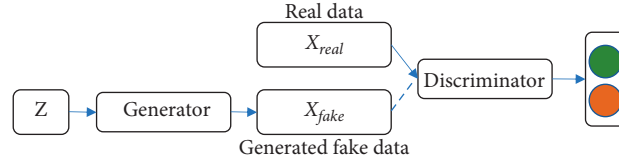
FIGURE 1: A counternetwork structure.

Through comparative analysis, it is found that the threshold denoising algorithm is the simplest and has the least amount of calculation in the image processing process, has the characteristics of high flexibility and strong denoising ability, and is the most widely used in practical applications. Therefore, this article chooses to use the threshold denoising algorithm to process the image.

3.1. Image Noise Reduction Preprocessing.

In a noisy image, the real signal of the image is usually a low-frequency or relatively stable signal, while the noise is generally concentrated in high-frequency signals. Wavelet transform can effectively concentrate energy and concentrate it in the low-frequency region. The Gaussian noise presents a Gaussian distribution, and the main signal is distributed in the high-frequency region. Therefore, the proposed algorithm can quantify the threshold of wavelet coefficients by introducing appropriate thresholds to eliminate noise.

Through the study of noisy images, one can find that, in real life, the low-frequency part of the image signal after wavelet transform is composed of the useful signal of the image, and the image noise often exists in the high-frequency part of the image [12].

Step 1: wavelet decomposition of the noisy signal: choose the appropriate wavelet and the corresponding decomposition layer. Figure 2 shows a wavelet decomposition diagram of a noisy image.

Step 2: perform threshold quantization processing on the high-frequency subbands after wavelet decomposition. Threshold processing of different scales is realized by the threshold, and the corresponding estimated wavelet coefficients are obtained.

Step 3: reconstruct the obtained signal through your transformation of wavelet coefficients. Finally, the denoised signal is obtained. That is, the wavelet coefficients of the low-frequency subband and other high-frequency subbands are quantized by the threshold, and the signal is reconstructed to obtain the latest estimated signal $\hat{f}(t)$.

The selection and quantification of the threshold and the classic threshold denoising function are as follows.

Threshold:

$$\lambda = \sigma * \sqrt{2 \log N}. \tag{3}$$

By measuring each detail signal, it can be filtered. Among them, σ is the standard deviation of the noise, and N is the length or scale of the signal. It is defined as follows:

$$\widehat{\omega}_{i,j} = \begin{cases} \omega_{i,j}, & |\omega_{i,j}| > \lambda, \\ 0, & |\omega_{i,j}| \leq \lambda. \end{cases} \tag{4}$$

Finally, the reconstructed data of the image can be obtained.

3.2. Noise Reduction (Detail Protection).

Combining the inherent visual characteristics of the human eye, this paper finds that the entropy value of the image can truly reflect the image signal [13–19], but in the actual computer process, the maximum entropy threshold is calculated in a relatively large amount. Therefore, in the process of noise reduction, after the graying of the image, because the image signal is mainly distributed with area 0 and region 1, the probability in regions 2 and 3 is usually set to 0, which greatly simplifies the calculation in the mathematical model.

Research found that, after adaptive decomposition, the edge signals in the high-frequency signal are effectively concentrated [20–28], while the high-frequency signal is mainly used to represent the edge signal of the image, which inevitably shows some noise signals and weak edge signals, if these signals are deleted, it will not affect the overall quality of the image, and how to effectively separate these signals from the edge signal will have a certain impact on the compression of the image.

By analyzing the edge characteristics of its image, we find out the characteristics:

(1) Along the gradient direction, the detection is carried out in a certain range, the maximum value is retained, and $M_S f(i, j) = 1$ can be obtained. If the nonmaximum value is deleted, $M_S f(i, j) = 0$ can be obtained.

(2) The edge signal is generally a point with a more intense grayscale change, i.e., the corresponding mode value is relatively large, while the corresponding mode value of the noise signal is relatively small.

Therefore, according to the distribution characteristics of the image signal, this paper classifies the image to achieve the compression processing of the image. The method is as follows:

$$W(i, j) = \begin{cases} W(i, j), & M_s f(i, j) > = \lambda, \\ 0, & M_s f(i, j) < \lambda. \end{cases} \tag{5}$$

Edge signal function: $M_s f(x, y) = s\sqrt{(\partial F/\partial x)^2 + (\partial F/\partial y)^2}$.
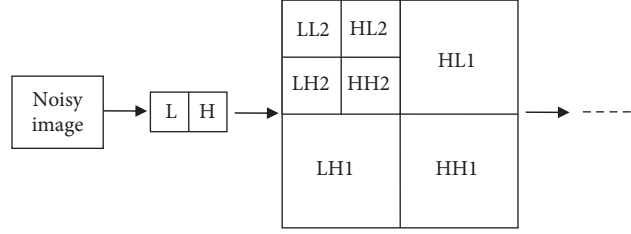
FIGURE 2: Wavelet decomposition diagram of the noisy image.

$$F_s(x, y) = \sum_{(i,j) \in N_s(x,y)} f(i, j) \phi_s(x - i, y - j),$$

$$\frac{\partial F_s(x, y)}{\partial x} = \sum_{(i,j) \in N_s(x,y)} f(i, j) \frac{\partial}{\partial x} \phi_s(x - i, y - j),$$

$$\frac{\partial F_s(x, y)}{\partial y} = \sum_{(i,j) \in N_s(x,y)} f(i, j) \frac{\partial}{\partial y} \phi_s(x - i, y - j),$$

$$\frac{\partial}{\partial x} \phi_s(x - i, y - j) = \frac{1}{s^3} \frac{\partial \phi(u, v)}{\partial u}\Big|_{(u,v)=(x-i/s, y-j/s)},$$

$$\frac{\partial}{\partial y} \phi_s(x - i, y - j) = \frac{1}{s^3} \frac{\partial \phi(u, v)}{\partial v}\Big|_{(u,v)=(x-i/s, y-j/s)}.$$

$$(6)$$

Among them, $0 \le x, y \le N$.

The edge signal in the high-frequency signal is detected, the edge signal is classified by the reference threshold, the edge signal of the image is extracted to enhance the effective signal in the high-frequency signal, and the negligible signal in the image is deleted.

At the same time, in order to avoid excessive denoising, a new threshold is proposed; that is, $\lambda = 2^a \sigma \sqrt{2 \log(N^2)}$.

Among them, a is the number of filtered stages, N^2 is the maximum scale size of the image signal, and σ is the noise standard estimation of the contourlet transform. This paper reduces the appearance of "overstrangling" phenomenon by combining the nature of the image signal and thus improves the accuracy of the denoising algorithm.

According to the inherent characteristics of human vision, the article classifies high-frequency signal information and combines with a new threshold denoising algorithm to effectively remove image noise. Through experiments, the algorithm in this paper can reduce the coding overhead of the denoising algorithm while ensuring the quality of the denoised picture.

By observing Figure 3, it is not difficult to find that this algorithm can effectively remove speckle noise. It can be seen that the processed image has enhanced the contrast of the image, but there are still blurry edges and loss of details.

3.3. Adaptive GAN to Improve Image Quality. Image super-resolution, referred to as super-resolution (SR), generally refers to the magnification resolution. For example, if you change the resolution from 256×256 to 512×512, the magnification scale is 2. Obviously, this is an ill-posed problem of making pixels out of nothing, and there is no unique solution. The image is superdivided, and the application scenarios are naturally wide. The general method is to take the low-resolution (LR) image as the input of the method and process it to obtain the high-resolution (HR) image. Nowadays, quite a lot of papers are self-made LR-HR image pairs as training sets. For example, the LR is obtained by downsampling the original image HR, and then the mapping learning from LR to HR is performed, but when applied to practice, is the relationship between LR and HR a self-righteous "downsampling" relationship? This is probably unknown and difficult to simulate.

Due to the weak correspondence between the random input and pictures, GAN-generated images are prone to misalignment. Also, because the discriminator judges the image as a whole, the generated pictures have stronger continuity and can generate clearer pictures.

When processing images, the pixel loss function similar to the mean square error cannot recover some details that are lost in the downsampling process. Minimizing the mean square error is by finding a reasonable pixel average value, but the processing result of this method is usually too smooth, which will lead to poor perception quality.

The GAN provides a very advanced learning architecture in generating very high-quality natural image research. The architecture of the GAN helps to facilitate the reconstruction of regions that move into the search space. This may contain images that are as real as photos, making the images obtained by the GAN closer to natural images.

Each iteration of the generator network includes three stages [29–31], which enable the generator to quickly and accurately locate the focus area in the image.

(1) *Variation:* Introduce a random noise signal z. In the new round of update process using the gradient penalty loss function of this paper, and for the generation of the network gradient penalty loss function is $L(G, D)$ calculated only for noise data, which makes the network according to the parameters updated.

Figure 3: Denoising effect diagram of the GAN.

$$L(G, D) = E_{z \sim P_z}[D(z)] - E_{z \sim P_z}[D(z)] - \lambda_{\mathrm{gp}} E_{z \sim P_z}\left[\left(\left\|\nabla_x D(z)\right\|_2 - 1\right)^2\right],$$

$$L_{\mathrm{MSE}}(G, D) = L(G, D) + \lambda_{\mathrm{MSE}} L_{\mathrm{MSE}}(G). \tag{7}$$

Among them, λ_{gp} is the weight parameter. P_z is the noise data distribution.

(2) Assess:

In order to learn the correct distribution of training samples and improve the stability of model training, this paper selects an evaluation mechanism that uses the loss function to select the best sample under the screening of the discriminator. The following is the evaluation mechanism:

The quality score of the generated samples and the diversity of the generated samples are the embodiment of the quantification of the adaptive score at this stage: $\mathbf{F} = \mathbf{F_q} + \gamma \mathbf{F_d}$.

$\mathbf{F_q}$ is the quality score, $\mathbf{F_d}$ is the diversity score, and γ is the ratio parameter of the two. Through the evaluation mechanism, the highest adaptability score is calculated, and the size of the highest score is calculated so that the global optimal score can be obtained. By analogy, until the end of the first generation to the last generation, the optimal value is obtained by comparison, and then the adaptive score is updated.

(3) Select:

According to the evaluation mechanism to select the optimal individual, the other low scores will gradually decay so that the generation network has a certain direction, and choose the optimal value of the individual.

4. Experimental Results and Analysis

In order to verify the validity of this method, verify that the algorithm can denoise while retaining image edge information. Since the GAN is a deep learning model, the game between the generation network and the confrontation network is used to improve the results of the generation network. In this paper, PyCharm is used to simulate the algorithm and to compare with the existing traditional algorithm [32] and literature algorithm [33] by processing the same noisy image.

4.1. Objective Indicators of Image Denoising. The experimental hardware environment is the processor Intel Corei7-7700, clocked at 3.60 GHz, memory 12 GB, NVIDIA TITAN Xp. The software environment is Windows 10, 64 bit, Python 3.6, and TensorFlow. In order to verify the denoising effect of the proposed algorithm on noisy images, the ImageNet image dataset is selected in the article and combined with the images in real life; 100×100, 128×128, and 256×256 images are selected. Tests were performed at Gaussian noise levels of 10, 20, 30, 40, and 50. On the experimental dataset, when the Gaussian noise standard deviation σ of the noise image is larger, the peak signal-to-noise ratio (PSNR) is smaller, and the image distortion is more. On the contrary, the smaller σ, the larger the PSNR and the better the denoising effect.

This paper verifies the fidelity effect of the algorithm by comparing the PSNR value of each algorithm. Table 1 is a comparison data table of the PSNR value obtained by different denoising algorithms through experiments.

It can be seen from Table 1 that, in the case of different noise standard deviations of the same noisy image, by analyzing the PSNR value, it can be seen that the data obtained by the algorithm in this paper are better than the literature algorithm and the traditional algorithm. The data of the algorithm in this paper are compared with those of the traditional algorithm, and the average PSNR value is 10.48 dB higher. Compared with the literature algorithm, the PSNR value of the algorithm in this paper is 6.56 dB higher, which effectively proves the image fidelity capability of the algorithm in this paper.

In order to verify the influence of the noise figure on the algorithm time, this article records the time consumed by the algorithm under the aforementioned environment and draws it into a graph, as shown in Figure 4.

By observing Figure 4, it is not difficult to find that, as the noise figure increases, the time consumed by each algorithm increases. Among them, the curve of the literature algorithm

TABLE 1: The PSNR value data table for the results obtained by different denoising algorithms.

σ	PSNR (dB)			
	Noisy images	Traditional denoise	Literature algorithms	Article algorithm
10	24.21	29.32	35.95	37.05
20	21.34	29.12	31.02	38.21
30	18.91	25.15	30.12	39.34
40	16.32	30.57	32.58	37.12
50	14.22	26.11	30.22	40.97
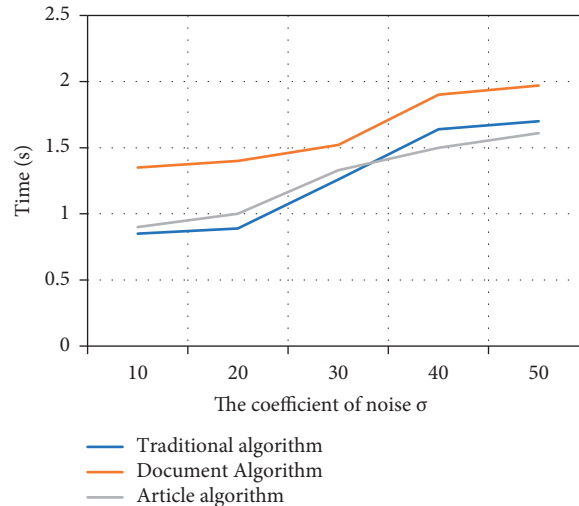
FIGURE 4: The influence curve of the noise figure on algorithm time.

consumes the most, both of which are above the time curve of the traditional algorithm and the article algorithm. The time curve obtained by the algorithm of the article is basically the same as the traditional algorithm.

4.2. Iterative Effect Comparison Experiment.
This article observes the PSNR value based on the ImageNet dataset and the CIFAR-10 dataset. We analyze the change data of PSNR as the number of iterations increases and express it through a graph. Figure 5 shows the PSNR curves of three different sizes of images as the number of iterations increases.

It is not difficult to find from the figure that the algorithm in this paper has a better denoising effect in different sizes. After the number of iterations reached 15 K, the growth rate of PSNR also slowed down, and finally, our network results reached a stable value.

4.3. Running Time.
In order to verify the algorithm, the experiment selects standard image library images and evaluates the average running time of the algorithm based on the same standard. The results are shown in Table 2.

It is not difficult to see from Table 2 that the algorithm in this paper is significantly better than the traditional algorithm in terms of speed. The reason is that although the computational complexity of the literature algorithm code has been significantly reduced after multiple optimizations, the attention-based denoising algorithm needs to extract images for iterative training models to achieve effective denoising, so the computational complexity is relatively high. However, the algorithm in this paper can significantly improve the operating efficiency by simply reducing it for the first time and then scoring and selecting the image. It can be seen from Table 2 that in terms of time, the time of the algorithm in this paper is basically the same as that of the traditional algorithm. These prove the feasibility of the algorithm.

4.4. Denoising Effect Comparison Experiment.
In order to verify the superiority of the algorithm in the image denoising effect, the pictures of the two actual scenes are selected for simulation experiments, and the denoising effect of each algorithm is shown, where Figures 6 and 7 show the original images.

By observing Figures 8–10, one can see that the image effect processed by the algorithm in this paper is significantly improved compared with the traditional algorithm and the literature algorithm. And the image details and image sharpness processed by the algorithm in this paper are better than the other two algorithms.

The proposed algorithm selects a complex image to perform related operations to verify the algorithm's processing effect on image details. Figure 7 is a high-altitude shooting with complex content. The effect is shown in the following.
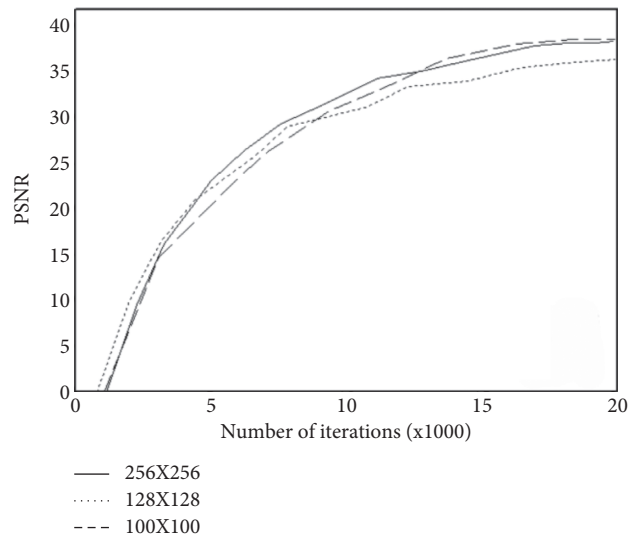
FIGURE 5: The PSNR curve of three different sizes of images as the number of iterations increases.

TABLE 2: Average running time of different algorithms.

Algorithm	Panda (s)	Duck (s)	Cliff (s)	Average running time (s)
Traditional algorithm	0.823	0.992	0.859	0.891333
Document algorithm	1.467	1.036	1.904	1.469
Article algorithm	0.993	0.838	0.882	0.904333

FIGURE 6: Original image.
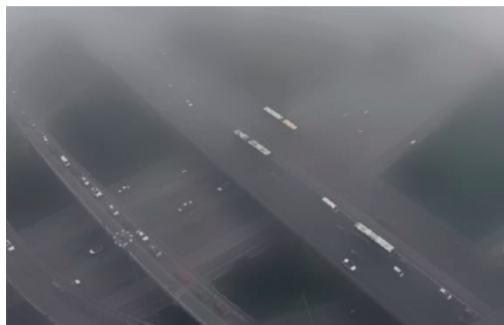
FIGURE 7: Original image.

Figure 8: Traditional algorithm.

Figure 9: Document algorithm.

Figure 10: Article algorithm.

Figure 11: Traditional algorithm.

FIGURE 12: Document algorithm.

FIGURE 13: Article algorithm.

By observing Figures 11–13, one can see that the algorithm in this paper compares the noise reduction simulation of the traditional algorithm and the literature algorithm. It can be seen that the algorithm can denoise the noisy image while still retaining the edge signal of the image and has a good performance. It has better visual effect.

5. Conclusion

This paper proposes a self-adjusting generative confrontation network image denoising algorithm. The algorithm includes noise reduction and the adaptive learning GAN model. It preprocesses the image first by combining image characteristics to make the target clear. Then, the algorithm in this paper uses the GAN model with adaptive learning to further denoise the image and then obtains the best denoising effect. The threshold value quoted in the process of noise reduction avoids the phenomenon of "overstrangling" and improves the effective signal in the high-frequency signal. In the GAN model, iterative updates are performed through operations such as mutation, evaluation, and selection. Self-learning of the generator can quickly and accurately locate the key areas in the image to obtain the optimal value. The results show that the proposed algorithm can effectively denoise while retaining the image signal, which is consistent with the expected effect. However, with the explosive growth of the amount of image data, how to optimize the algorithm to improve the efficiency of image denoising execution is a problem that needs further research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

The authors contributed equally to the manuscript. All authors read and approved the final manuscript.

Acknowledgments

References

[1] K. Shi, "Image denoising by nonlinear nonlocal diffusion equations," *Journal of Computational and Applied Mathematics*, vol. 395, no. 7, Article ID 113605, 2021.

[2] W.-H. Lee, M. Ozger, U. Challita, and K. W. Sung, "Noise learning-based denoising autoencoder," *IEEE Communications Letters*, vol. 25, no. 9, pp. 2983–2987, 2021.

[3] E. T. Tan, S. C. Queler, B. Lin et al., "Improved nerve conspicuity with water-weighting and denoising in two-point

Dixon magnetic resonance neurography," *Magnetic Resonance Imaging*, vol. 79, pp. 103–111, 2021.

[4] Y. Zhou, R. Chen, Y. Zhao, X. Ai, and G. Zhou, "Point cloud denoising using non-local collaborative projections," *Pattern Recognition*, vol. 120, no. 81, pp. 108–128, 2021.

[5] J. Park, C.-K. Kang, and Y. Lee, "Quantitative evaluation of the image quality using the fast nonlocal means denoising approach in diffusion-weighted magnetic resonance imaging with high b-value," *Journal of the Korean Physical Society*, vol. 78, no. 3, pp. 244–250, 2021.

[6] Z. Wan, "Infrared image enhancement algorithm using wavelet contourlet coefficients dependency," *Journal of Graphics*, vol. 35, no. 01, pp. 93–99, 2014.

[7] W. Niu, J. Meng, Ze Wang, and K. Cui, "Image denoising based on adaptive contraction function contourlet transform," *Journal of Graphics*, vol. 36, no. 04, pp. 593–602, 2015.

[8] J. Luo and J. Huang, "Generative adversarial network: an overview," *Chinese Journal of Scientific Instrument*, vol. 40, no. 3, pp. 74–84, 2019.

[9] Z. Li, Z. Shi, and C. Zhang, "Hybrid generative/discriminative model for automatic image annotation," *Journal of Image and Graphics*, vol. 20, no. 5, pp. 687–699, 2015.

[10] Y. Cao, L. Jia, and Y. Chen, "Review of computer vision based on generative adversarial networks," *Journal of Image and Graphics*, vol. 23, no. 10, pp. 1433–1449, 2018.

[11] K. Wang, C. Gou, Y. Duan, L. Yilun, X. H. Zheng, and F. Y. Wang, "Generative adversarial networks: the state of the art and beyond," *Acta Automatica Sinica*, vol. 43, no. 3, pp. 321–332, 2017.

[12] K. Asem, R. R. Abd, J. B. H. Shaiful, and A. R. Syed, "Additive noise reduction in natural images using second generation wavelet transform hidden Markov models," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 11, no. 3, pp. 339–347, 2016.

[13] S. Ye, "Image enhancement algorithm combining with threshold de-noising and edge optimization," *Journal of Graphics*, vol. 35, no. 04, pp. 571–576, 2014.

[14] G. Chen, F. Zhu, and P. A. Heng, "An efficient statistical method for image noise level estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, December 2015.

[15] X.-H. Wang, Yi-H. Zhu, F. Lv, X. Su, and C. M. Song, "Cauchy distribution NSST-HMT model and its applications in image denoising," *Chinese Journal of Computers*, vol. 41, no. 11, pp. 2496–2508, 2018.

[16] H.-Y. Jin, J. I. A. O. Li-Cheng, and F. Liu, "SAR image denoising based on curvelet domain hidden Markov tree models," *Chinese Journal of Computers*, vol. 1, no. 03, pp. 491–497, 2007.

[17] C. Yang, Y. Luo, and S. Yang, "Hybrid linear model based image denoising," *Chinese Journal of Computers*, vol. 32, no. 11, pp. 2260–2264, 2009.

[18] L I N. Zhou-Chen, "AN anisotropic diffusion equation that can remove noise and keep naturalness," *Chinese Journal of Computers*, vol. 1, no. 11, pp. 1133–1137, 1999.

[19] B. Jing, B. Hou, S. Wang, and L. Jiao, "SAR image denoising based on lifting directionlet domain Gaussian scale mixtures model," *Chinese Journal of Computers*, vol. 1, no. 07, pp. 1234–1241, 2008.

[20] R. Ramlee, "Blur removal in natural digital images using self-reference generative networks," *JTEC*, vol. 13, no. No.3, pp. 61–65, 2021.

[21] K. Liu, R. Ma, and Y. Pang, "A detail enhancement and denoising algorithm of high dynamic range infrared image based on double guided image filter," *Journal of Graphics*, vol. 39, no. 06, pp. 1048–1054, 2018.

[22] W.-hao Wang, Y. A. N. Yun-yang, M.-xin Jiang, S.-bing Gao, and Y.-tao Yu, "Image denoising algorithm based on noise detection and dynamic Window," *Journal of Graphics*, vol. 40, no. 01, pp. 111–116, 2019.

[23] G. B. Riya, B. Gupta, and S. S. Lamba, "An efficient anisotropic diffusion model for image denoising with edge preservation," *Computers & Mathematics with Applications*, vol. 93, no. 4, pp. 106–119, 2021.

[24] Y. Quilfen and B. Chapron, "On denoising satellite altimeter measurements for high-resolution geophysical signal analysis," *Advances in Space Research*, vol. 68, no. 2, pp. 875–891, 2021.

[25] N. Salamat, M. Missen, and V. Prasath, "Recent developments in computational color image denoising with PDEs to deep learning: a review," *Artificial Intelligence Review*, vol. 54, no. 8, pp. 1–32, 2021.

[26] J. Yu and G. Liu, "Extracting and inserting knowledge into stacked denoising auto-encoders," *Neural Networks*, vol. 137, pp. 31–42, 2021.

[27] C. Hu, C. Zhou, and S. Wang, "Research based on mathematics morphology image chirp method," *Journal of Engineering Graphics*, vol. 1, no. 02, pp. 116–119, 2004.

[28] W. Gao, L. I. Zheng, and K. A. N. G. Qian, "A hybrid mesh denoising algorithm based on the mean normal filter," *Journal of Engineering Graphics*, vol. 32, no. 04, pp. 84–89, 2011.

[29] H. Wang, Y. A. N. G. Xiao, Y. Jiang, and Z. Wang, "Image denoising algorithm based on multi-channel GAN," *Journal on Communications*, vol. 42, no. 03, pp. 229–237, 2021.

[30] Y. He and M. Yang, "OCT image denoising method based on generative confrontation network," *Modern Computer*, vol. 1, no. 12, pp. 87–91, 2021.

[31] X. Cai, S. Zhang, Q. Chen, Y. Chen, and M. Wu, "Structure preservation generative adversarial network for noise reduction in SD-OCT images," *Journal of Computer-Aided Design & Computer Graphics*, vol. 32, no. 05, pp. 751–758, 2020.

[32] ZW.L. I Wang and Zi Shao, "Adaptive fractal-wavelet image denoising base on multivariate statistical model," *Chinese Journal of Computers*, vol. 37, no. 06, pp. 1380–1389, 2014.

[33] Q. Lyu, M. Guo, and M. Ma, "Boosting attention fusion generative adversarial network for image denoising," *Neural Computing & Applications*, vol. 33, no. 10, pp. 4833–4847, 2021.

Research Article

Research on Robot Fuzzy Neural Network Motion System Based on Artificial Intelligence

Jie Hu (ID)

Suzhou University, Institute of Physical Education, Suzhou 234000, Anhui, China

Correspondence should be addressed to Jie Hu; szxyhujie@ahszu.edu.cn

An intelligent controller based on a self-learning interval type-II fuzzy neural network is proposed to make the motion controller of the industrial intelligent robot with good adaptability. This controller has a parallel structure and contains an interval type-II fuzzy neural network and a conventional PD controller. For the design of the interval type-II fuzzy neural network, the interval type-II fuzzy set is established using the slave design method. In the design process of the interval type-II fuzzy set of the front piece, a dual sequence symmetric trapezoidal subordinate function arrangement method is proposed, which makes the self-learning law and stability analysis of the system in an analytic form and facilitates the implementation of the algorithm in hardware. In the design of the neural network self-learning law, a parametric self-learning algorithm based on sliding mode control theory is established to adjust the structural parameters of the interval type-II fuzzy neural network online, and the stability of the system is proved by using Lyapunov's stability theorem. Three sets of validation simulation experiments are given in conjunction with the trajectory tracking problem of the Delta parallel robot. The simulation results show that, in the presence of system uncertainty, the intelligent controller based on interval self-learning interval type-II fuzzy neural network can significantly improve the trajectory tracking accuracy and robustness of the system and make the control system highly adaptable to the environment. Experiments of intelligent control system based on self-learning interval type-II fuzzy neural network and experiments of reusable particle swarm optimal motion planning method are designed, and the effectiveness of the intelligent control system and motion planning method is verified on the experimental platform. The experimental results show that the intelligent control system based on the self-learning interval type-II fuzzy neural network can effectively improve the accuracy and stability of robot trajectory tracking control, and the reusable particle swarm optimal motion planning method can quickly solve the robot motion planning problem with complex constraints online.

1. Introduction

Artificial intelligence is both the background of the times and a prerequisite for machines to have intelligence and sociality. Existing studies on human-machine social interaction only consider the shallow expression of human-human social or human-machine interaction, without considering the trends and design requirements of the deeper changes in the human-machine relationship in the upcoming AI era. In the past, machines were used as tools or mediums for human-human interaction, but with the further enhancement of machine intelligence, direct human-machine interaction has gradually evolved [1]. Through the increasingly in-depth human-machine interaction, intelligent machines have functional, emotional, and social effects on people, which leads to the new topic of human-machine social interaction in the era of artificial intelligence [2]. In the context of the new social environment and the development of new human-machine social interaction, we start from the background and the current situation of the study, use literature review and typical example analysis in the early stage of the study to understand the current research process of artificial intelligence, social interaction, and human-machine interaction, explore the part that can be further developed based on the shortage of existing research, and propose the purpose and significance of this study. When IT2FNN is connected in parallel to both ends of the PD controller, the self-learning process of IT2FNN's

structural parameters begins. After the learning time of about 1.7 seconds is over, the output signal of IT2FNN replaces the PD controller as the main output source of the control signal. The result is the same. The above simulation results clearly show the influence of the learning process of SLIT2FNN on the output of the controller and the improvement of control accuracy after adopting the parallel intelligent control structure in the second working cycle.

In practical applications, one of the most frequently encountered situations for industrial intelligent robots is trajectory tracking control [3]. If a satisfactory control effect is to be obtained for an industrial robot, there are many problems to be solved in the design of the control system, such as the nonlinearity of the robot body structure, the uncertainty of modeled and unmodeled dynamics, and the diversity of operating conditions, in addition to its dynamics optimization design. In the last decade, many related research works have been done by scholars on the trajectory tracking control problem of parallel robots. Some of them are based on traditional control methods for trajectory tracking control problems: some studies [4, 5] are based on feedback linearization control methods; others [5, 6] are based on adaptive theory, in addition to studies based on variable structure control theory. However, most of the controllers based on traditional control methods need to provide an accurate mathematical model of the controlled object to obtain satisfactory control results [7]. However, obtaining a complete dynamical model of a parallel robot is very difficult and sometimes impossible. Moreover, even if accurate robot system parameters can be obtained experimentally, it is equally difficult to derive a dynamic model of the robot based on these parameters, especially for some parallel robot systems with complex motion mechanisms.

Taking the Delta parallel robot as the object, the forward and inverse kinematics models of the robot are established, and the dynamics model of the robot is established by using the virtual work principal method. On this basis, a robot body mechanism parameter optimization method oriented to the control task requirements is proposed. Firstly, the mechanism parameter optimization algorithm considering anisotropic dynamics characteristics is established, and the optimization models are built with the full-domain drive joint torque index and the full-domain drive joint torque fluctuation index, respectively, and the optimization problems are solved to verify the effectiveness of the algorithm. Then a multiobjective dynamics optimization model based on the full-domain performance index is established, and the multiobjective optimization problem is solved using a multiobjective genetic algorithm, and the computational efficiency and effectiveness of the algorithm are verified. The above model and optimization algorithm of the Delta parallel robot are established using the idea of building blocks to make the robot system simulation platform universal. A parallel intelligent control system based on a self-learning interval type-II fuzzy neural network is proposed and used to solve the trajectory tracking control problem of the Delta parallel robot. This controller is a parallel structure: it contains an interval type-II fuzzy neural network and a conventional PD controller. The IT2FS is designed using a

subordinate design approach, based on which a dual sequence symmetric trapezoidal subordinate function arrangement is proposed to give an analytic form to the interval type-II subordinate function for designing the self-learning law and stability analysis. In the design of the neural network self-learning law, a parametric self-learning algorithm based on the SMC theory is proposed to adjust the structural parameters of the interval type-two fuzzy neural network online. The stability of the proposed intelligent control system is proved using Lyapunov's stability theorem. The performance of the established intelligent controller is verified through simulation experiments.

2. Related Work

In terms of isotropic optimization problems in the full working domain, Narayan et al. proposed a multiobjective optimization method for the design of structural parameters of a four-degree-of-freedom parallel robot for the robot isotropic design problem with task requirements, using robot sensitivity analysis and standardized workspace volume as objective functions [8]. Rao et al. studied the scale synthesis problem of the Hexaslides robot and gave a method to select the structural parameters of the Hexaslides robot for the machining field [9]. Mirzaey et al. investigated the kinematic optimization problem of a two-degree-of-freedom parallel robot arm using a stepwise optimization method. The number of design variables was reduced by optimizing the parameter relations with an analytic form in the first step [10]. The second step reduces the optimization variables of the optimization problem to one utilizing the given global optimization index and thus solves the kinematic optimization problem of a parallel robotic arm [11]. Li et al. proposed a dynamic isotropic optimization method for the full working domain of a parallel robot, giving a dynamic optimization objective function based on the robot inertia matrix, and demonstrated the validity of its objective function through an optimal design problem of a three-degree-of-freedom tandem robot and a parallel robot [12].

Since we need to obtain a controller with high nonlinear characteristics and smoother control performance, the IT2FLC obtained by the two-terminal fuzzy method is the structure of the type 2 fuzzy logic controller that best meets the expectations of our control system characteristics. Vision servo is based on the visual information provided by the camera to control the robot. Depending on the feedback error used to calculate the control law, the vision servo is divided into position-based, image-based, and motion-based control systems [13]. In the position-based control system, the feedback error is calculated in a three-dimensional Cartesian coordinate space; in the image-based control system, the feedback error is calculated only from the different pixel points of the current image and the reference image; both systems are model based and require rigorous calibration of the camera to determine the exact geometric relationships between the camera and the robot's end-effector and then use these geometric relationships to analyze the feature points in the image [14]. These geometric relationships are then used to analyze the feature points in the

image, which are usually selected by humans. In contrast, motion-based control systems are model-free and can calculate feedback errors without any a priori using of frame difference methods, optical flow methods, etc. With the development of technologies such as deep learning and the advent of the era of artificial intelligence, the connection between machines and people is getting closer and closer, and artificial intelligence with human-machine dialogue function is gradually social role-playing, and the human-machine relationship is changing accordingly [15]. The research finds that existing studies on social interaction only consider human-human socialization and do not consider the changes in the human-machine relationship in the coming AI era [16]. To verify whether the motion mode of the real machine is consistent with the motion gait based on the Hop oscillation model proposed in this paper, the wake gait motion of the robot dog can be obtained through the video screenshots during the experiment. This study explores the changes in human-robot social relationships, the deep-seated needs, and ultimate purpose of human-robot social interaction and analyzes how evolving human-like intelligent robots can meet the dynamic deep-seated emotional needs of humans to adapt to human needs.

The communication between man and machine will rise from mechanical external interaction to an emotional level. All specific operational devices will be naturally integrated into the overall information infrastructure. In their place, there will be a wide variety of ubiquitous sensors. The integration of AI, cloud computing, and increasingly intelligent machines will generate data for use in human-computer interaction systems. The study of social interaction has the value of studying and predicting future trends in the evolution of human-machine relationships. Machines gradually take on the attributes of social roles and, on the one hand, have a social connection with people. From another perspective, the advancement of human-like intelligence such as artificial intelligence, which is mainly driven technology, has replaced some of the jobs originally performed by humans, thus causing an identity crisis and technological anxiety in human values. The development of artificial intelligence and human-like emotions, which are mainly emotionally dependent, will have a more profound impact on human society, especially in terms of human identity and emotional dependence in social relations.

3. Fuzzy Neural Network Motion System for Robot with Artificial Intelligence

3.1. Robot Motion System Design. To realize the robot from the virtual prototype model to the real model, the first step is to draw the vector drawing for the real model and to mark the drawing size on the vector drawing according to the size of the virtual prototype; the common software to draw the structure drawing is Autodesk Computer-Aided Design (AutoCAD) and CorelDRAW. In this paper, we chose AutoCAD as the drawing tool.

The body structure of the robot is three-dimensional. In this paper, the length, width, and depth of each part need to be marked in the drawing design so that the robot body can

be assembled by printing out the parts with a three-dimensional structure [17]. The body structure of the quadruped robot was designed concerning the mortise and tenon structure of ancient Chinese architecture, and the torso of the robot was designed to be fixed without screws, but only by inserting the mortise and tenon structure. After the design of the drawing, the drawing was processed by laser engraving, and then the dimensions of the drawing were verified by assembling, and the unreasonable structural parts were modified.

The robot's control method is divided into two types: external command control and automatic adjustment control of the robot's decision system [18]. In external command control, the robot needs to have a signal receiver that can communicate with the outside, and in decision control, the robot needs to be able to sense the surrounding road conditions, audio, and video signals, etc. After receiving the signals, the robot's brain needs to perform a series of operations, such as data calculation and task scheduling, and then control the leg servos to make corresponding movements. The SLIT2FNN controller is in the environment dynamics learning process. The neural network adjusts its own structural parameter values through the parameter self-learning algorithm. At this time, the control signal is mainly provided by the PD controller in the parallel control structure. After the neural network self-learning process is over, IT2FNN replaced the traditional PD controller as the main control signal output in the parallel control structure. At the same time, it can be seen from the figure that the tracking accuracy of the robot has been significantly improved. Therefore, the hardware structure of a robot is like the von Neumann architecture of a computer, which consists of five major parts, each of which cooperates to realize the robot's functions. The robot hardware architecture is shown in Figure 1, where the communication module, the video perception module, and the pose perception module are the input devices of the robot, the servo controlled by PWM (pulse width modulation) signal is the output device of the robot, and the middle part is the operator and controller of the robot.

The robot system framework is a typical hierarchical multiloop control structure, in which the robot system is divided into different layers from top to bottom, and each layer will have multiple tasks to be executed collaboratively. The first three layers are embedded inside the robot, and the last layer is the interaction layer between external users and the robot.

The main function of the driver layer is to realize the cooperation between the robot hardware and software. The sensor interface is responsible for receiving digital signals from the sensors installed on the robot body, such as receiving video stream signals from the camera, while the main command of the actuator interface is to send commands to the actuator, sending the system control signals to the actuator and controlling the robot's servo to complete the motion function [19]. Firstly, the mechanism parameter optimization algorithm considering the anisotropic dynamic characteristics is established, and the optimization model is established with the global driving joint torque index and the
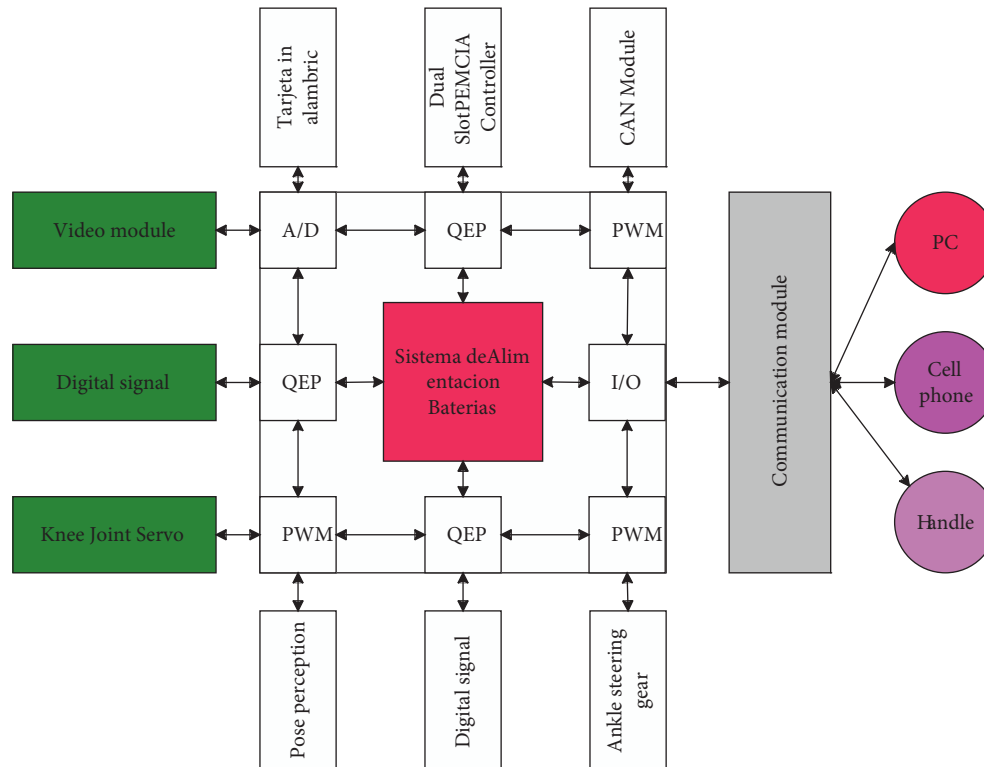
FIGURE 1: Hardware structure of the robot.

global driving joint torque fluctuation index, respectively, and the optimization problem is solved to verify the effectiveness of the algorithm. The main command of the actuator interface is to send commands to the actuator, send the system control signals to the actuator, control the robot's servo to complete the motion function, and receive the feedback signals from the actuator to sense whether the actual motion of the robot is the same as the control command. The second layer of the multitasking platform layer mainly implements the fusion of sensor data, control of multiple servos, image processing, and task scheduling functions, and the platform layer requires the driver layer and the algorithm layer together to complete the task. The robot system needs to fuse these sensor data and then realize the processing of different sensor data through the corresponding algorithms in the upper algorithm layer. The servo control interface mainly sends the control signal of the servo to the actuator and realizes the control of a certain servo or the parallel control of the whole-body servo according to the task, and the image processing interface accepts the video stream data sent by the camera sensor, and then the image processing interface of the algorithm layer performs calculations, as shown in Figure 2.

The specific manifestation of social relations is the human-computer relationship. Human-computer relations place more emphasis on the individual characteristics of people who influence each other, while social relations refer to the common aspects it contains. Social relations refer to the interpersonal relations formed by people in the process of production and living together. Social relations are mainly

divided into blood relations, georelations, and karma relations. Blood relations are rooted in or kinship ties and are the first social relations formed by human beings. They constitute social groups such as families and households. Georelationships are based on the spatial and geographical relationships of people. Karmic relations are formed based on the extensive social division of labor and promote social development. Types of interpersonal relations refer to the distinction of types of interpersonal relations. According to Thibault's view of interpersonal interdependence, he classifies interpersonal relationships as minimal interdependence, asymmetric interdependence, reactive interdependence, and mutual interdependence. Minimal interdependence means that the interacting parties act according to their intentions. Asymmetric dependence means that one party's actions are based on the actions of the other party, but with opposite intentions. Reactive interdependence means that one party acts in full accordance with the other party's intentions. Mutual dependence is the most equal and stable form of interaction; e.g., it can be classified as emotional or instrumental interpersonal relationships according to needs.

In terms of the relationship and process of human-computer interaction, the user (human) is constantly changing during the use of the machine. Human behavior and emotions are not stable; people are not stable; people are rather a role that is constantly changing with the use of intelligent machines. This calculation example only considers the situation that there are 1 and 2 obstacles in the workspace, each database contains 180 sets of data, and the
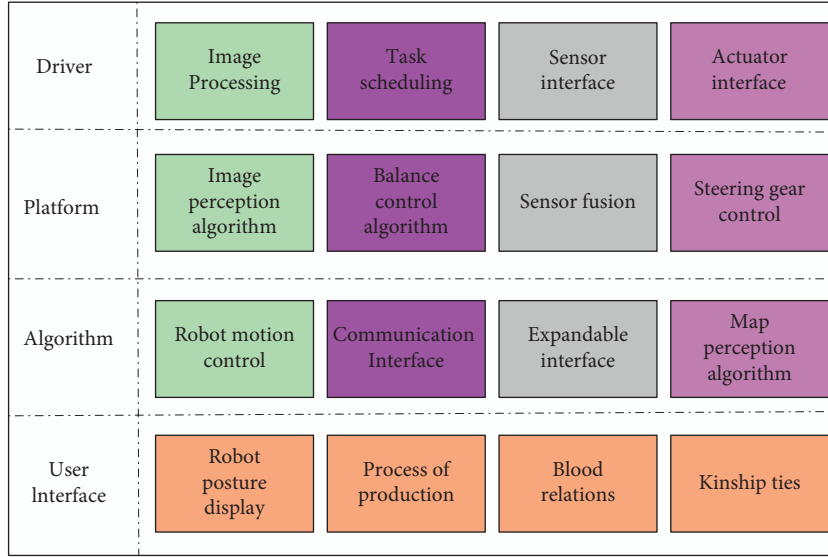
FIGURE 2: Hierarchical circuit diagram of the robot system.

setting of the initial motion trajectory of the robot is the same as the previous one. The dynamic performance of the social role of the machine is that the process of social role formation is affected by other factors, such as the degree of the flow of human-computer interaction, the degree of human involvement in the interaction, and so on. These factors or conditions are flexible and uncontrollable. In addition, the effect of the social role that can be achieved varies depending on the degree of machine intelligence. The social role properties of machines are not invariable; a machine is a social role only when it interacts with and influences people; in other words, active and interacting machine intelligence is a machine social role. With the support of big data and the progress of deep algorithms and other technologies, human beings have entered the era of artificial intelligence. Modern human-robot interaction is based on machines running according to programs written in advance by humans. In essence, human intentions reflected in the behavior of all robots. Intelligent machines take certain actions based on preprogrammed instructions from human programmers. Even intelligent robots make certain decisions on their own, based on algorithmic judgment. However, with machines having dialogue and behavioral feedback systems that are highly close to humans, close to real people's natural language understanding and output and behavioral feedback, human-robot social interaction arises along with it. Intelligent robots are a completely new product born in recent decades, and human-machine relationships and human-machine interactions are also a completely new subject. The construction and understanding of human-machine social interactions are both to help robots integrate into human and social systems and to build systems for good human-machine interactions. Machines becoming social actors is a prerequisite for human-robot social interaction. Intelligent machines with dialogue and behavioral feedback systems that are highly like those of humans and capable of adaptive autonomous learning become social actors with certain norms and dynamically changing behavior patterns when they participate in human-robot social interaction as subjects.

3.2. *Fuzzy Neural Network Algorithm Design.* Designing a type-one fuzzy logic system involves making many choices for each segment. We need to choose the type of fuzzification, the form of the affiliation function, the parameters of the affiliation function, synthesis, implication, and the type of demulsified, etc. The different choices for each link represent the different mathematical methods involved. For example, the firing level of a rule R^n can be written in the following form if it is calculated using the minimum t-parametrization. Although the EFPF-PSO algorithm has a chance to obtain an optimized trajectory with a smaller fitness value, not every calculation will converge to the minimum fitness value, and its standard deviation is relatively large. However, the results obtained by the RPSOMP method are more stable, the standard deviation is small, and the convergence speed is faster.

$$\mu_{R^n}(x) = \arg\max\left\{\mu_{A_i^n}\left(x_i^2\right)\right\}. \tag{1}$$

The output fuzzy set can be obtained by calculating the ignition level and the fuzzy rule posterior fuzzy set using the t-parametric operator. The final output fuzzy set B can be obtained by combining all the output fuzzy sets by t-residual van. To obtain the output clear value y, the output fuzzy set B needs to be defuzzied using a fuzzifier, and if the center of gravity fuzzifier is chosen, the output of the fuzzy logic system can be expressed as

$$y = \frac{\sum_{i=1}^{n} y_i \mu_B(y_i)}{\sum_{i=1}^{n} \mu_B(y_i)^2}. \tag{2}$$

The high computational effort and complex design process of general type-two fuzzy logic systems constrain

their application in control systems. In recent years, researchers in controller design have turned their attention to interval type-two fuzzy logic systems. The interval type-two fuzzy set A can be expressed as

$$A = \int_{x \in X} \int_{u \in J_x} (x, u).$$ (3)

When $f_x(\mu) = 1$, then the sub-subordinate function is an interval set, and if for both it holds, then the interval type-2 affiliation function (IT2MF, Interval Type-2 Membership Function) is obtained. The interval type-2 sub-subordinate function reflects the consistent uncertainty at the primary affiliation of x. Compared with the general type-two fuzzy logic system, the interval type-two fuzzy logic system greatly reduces the computational effort. Since interval type-two fuzzy sets also belong to type-two fuzzy sets, we can obtain

$$J_x = \{\mu \in [0, 1] | \mu_A(x, u) = 1\}.$$ (4)

In this paper, the center of sets type reducer is used to calculate the reduced type of the obtained ignition output set. Since both the antecedent and the consequent parts are interval type-two fuzzy sets, the obtained reduced typeset is still an interval set. The procedure of degenerate calculation: firstly, the output space Y is discretized into an appropriate number of points, and the center of gravity of the interval type-II fuzzy set of the posterior is calculated by the following equation:

$$C_{B_j^n} = \frac{\int_{\theta \in J_{y1}} \int_{u \in J_x} (x, u)}{\sum_{i=1}^{n} y_i \mu_B (y_i^2) \sum_{i=1}^{n} \mu_B (y_i^2)^2}.$$ (5)

To achieve the landing of the robot from the virtual prototype model to the real machine model, the first step that needs to be done is to draw a vector diagram for the real machine model. According to the size of the virtual prototype, the drawing size is marked on the vector diagram, and the commonly used drawing structure software for the drawing includes AutoCAD (Autodesk Computer-Aided Design) and CorelDRAW. In this article, we chose AutoCAD as the drawing tool. From equation (5), the center of gravity of the posterior interval type-II fuzzy set can be calculated in advance outside the control loop, which can reduce a large amount of computational burden for the control chip. The equation for the center-drop type of the set is

$$Y_{\cos} = \frac{\sum_{j=1}^{J} f^j}{\sum_{j=1}^{J} f^j y^j}.$$ (6)

Therefore, to calculate the center of the interval set $Y_{\cos}$, only its left endpoint y_l and right endpoint y_r need to be calculated. The calculation of the left and right endpoints is given by the following equation:

$$y_l = \frac{\sum_{j=1}^{L} f^j y^j - \sum_{j=L+1}^{J} f^j y^j}{\sum_{j=1}^{R} f^j},$$ (7)

$$y_r = \frac{\sum_{j=1}^{L} f^j y^j - \sum_{j=L+1}^{R} f^j y^j}{\sum_{j=1}^{J} f^j}.$$ (8)

The values of L and R in equations (7) and (8) are the values of the transition points needed to calculate the endpoint values, and the values of the left and right endpoints need to be calculated using the Karnik–Mendel (KM) algorithm. The design goal of the controller is to apply it to the trajectory tracking control of the Delta parallel robot and to provide a basis for subsequent research. The trajectory tracking control system of the Delta parallel robot uses three identical fuzzy logic controllers to control each of the robot's three servo motors. The demanded trajectory of the TCP is converted into the demanded angle of the drive joint by the kinematic inverse solution calculation and input to the demand trajectory of the TCP is converted into the demand angle of the drive joint by the kinematic inverse solution and input to the controller [20]. The input to the controller is the difference between the demand angle of the drive joint and the actual angle and its derivative. The T1FLC, which is the benchmark for the IT2FLC design, uses a type-one fuzzy set of trapezoidal affiliation functions as the antecedent and a type-one fuzzy set of triangular affiliation functions as the precedent.

The terms positive, negative, and zero used in the above fuzzy rules are expressed in the form of a type-one fuzzy affiliation function. The initial parameters of the type-one fuzzy set in T1FLC are empirically given artificially and adjusted to achieve a stable and continuous initial control performance of the robot system. The initial design of the input and output fuzzy sets applied to the trajectory tracking control of drive joint 1 is shown in Figure 3. The initial design structure of T1FLC for drive joint 2 and drive joint 3 is the same as that of drive joint 1.

Three types of IT2FLC are obtained using these three fuzzification methods, and these three types of IT2FLC are applied to the trajectory tracking control task of the Delta robot so that the TCP of the robot moves along the trajectory given by the following equation:

$$\begin{cases} x = 0.1 \cos(0.5\pi t + \pi) + 0.2, \\ y = 0.1 \sin(0.5\pi t + \pi), \\ z = 0.2547. \end{cases}$$ (9)

The fuzzy degree $\alpha \in [0, 1]$, which takes a value every 0.1, can construct one IT2FLC for each fuzzy degree value, and five trajectory tracking control simulations are performed for each constructed IT2FLC to reduce the statistical error of the simulation results. The root means square error (RMSE) values of the three joints obtained from the simulation are plotted as shown in Figure 4. It can be observed from the figure that the RMSE value increases with the increase of fuzziness, and the RMSE value obtained by the two-end fuzzy method is larger than the other two methods. The

reason for this phenomenon can be explained by Wu's work, when using the two-end fuzzy method, the FOU area produced by the same fuzzy degree is larger, and the larger FOU gives the fuzzy logic controller a larger damping characteristic. Robot control methods are divided into external command control and robot decision-making system automatic adjustment control. In the external command control, the robot needs to have a signal receiver that can communicate with the outside. The decision-making control requires the robot to be able to perceive the surrounding road conditions, audio, and video signal, etc. Increasing the damping characteristics of the controller can lead to more stable control performance, but it also causes a decrease in control accuracy. Therefore, the RMSE values obtained by using the two-terminal fuzzy method are larger than those obtained by the other two methods. The solution to the control accuracy degradation problem will be presented in the next section. In summary, since we need to obtain a controller with high nonlinear characteristics and smoother control performance, the IT2FLC obtained by the two-end fuzzy method is the structure of the type-two fuzzy logic controller that best meets the expectations of our control system characteristics.

The platform layer needs the driver layer and the algorithm layer to work together to complete the task. First, the platform layer receives the sensor signal of the driver layer. The quadruped robot has Bluetooth sensors and balance sensors on the body. The robot is always receiving multiple sensors during its movement. Signal, the robot system needs to fuse these sensor data and then realize the processing of different sensor data through the corresponding algorithm in the upper algorithm layer. The steering gear control interface is mainly to send the steering gear control signal to the actuator. This summary mainly tests the static gait walk state, as well as the dynamic gait trot state and backward motion and left-to-right turning motion of the robot dog proposed in this paper. The test environment was carried out on a smooth road to verify whether the motion pattern of the real machine and the motion gait based on Hop's oscillation model proposed in this paper are consistent, and the wake gait motion of the robot dog can be obtained through the video screenshots during the experiment. The robot dog can achieve a smooth and straight line, and the robot's walk state motion period T is 0.8 s, the step length of one cycle is 0.12 m, and the movement speed is 0.15 m/s through the test.

4. Analysis of Results

4.1. System Performance Results. A stable and effective interval type-II fuzzy logic control system can be designed using the dependent design method. Three simulation examples are used to investigate the control performance of the proposed SLIT2FNN controller. For comparative study, the conventional controller and the SLIT2FNN controller are connected separately to the Delta robot for trajectory tracking simulation experiments. The control goal is to make the TCP of the Delta parallel robot move along the given trajectory with as low as possible position and velocity errors

and achieve high trajectory tracking accuracy regardless of the internal and external uncertainties. The RMSE values of the angles of the three drive joints and the positions of the TCP in the workspace were recorded as the indexes for evaluating the control performance after the TCP of the robot ran along the trajectory given by equation (9) for one working cycle (8 seconds). The trajectory tracking errors of the drive joints and the TCP of the SLIT2FNN controller and the conventional PD controller are shown in Figure 5.

As we can observe in Figure 5, at the beginning of the simulation (about 1.7 s), the SLIT2FNN controller is in the process of learning the environmental dynamics, and the neural network adjusts its structural parameter values through the parameter self-learning algorithm, while the control signal is mainly provided by the PD controller in the parallel control structure. At the end of the neural network self-learning process, the IT2FNN replaces the conventional PD controller as the main control signal output in the parallel control structure, while the tracking accuracy of the robot is significantly improved as can be seen in the figure. The social interaction between man and machine will follow. Intelligent robots are a completely new product born in recent decades. Human-machine relationship and human-computer interaction are also a brand-new subject. The construction and understanding of human-computer social interaction is to help robots integrate into human and social systems.

A second simulation experiment is given to demonstrate more intuitively the learning process of the SLIT2FNN controller and the effect of parameter self-learning on the control performance. In this simulation, the same robot trajectory is used as in the first simulation, but the robot is required to run along this trajectory for two consecutive working cycles. In the first work cycle (8 seconds), only the conventional PD controller is involved in controlling the trajectory tracking of the robot. In the second work cycle, IT2FNN is connected in parallel to both sides of the PD controller to form a parallel control structure to complete the remaining trajectory tracking process. Figure 6 gives the output signals of IT2FNN and PD controller during this simulation experiment. It can be observed that when the IT2FNN is connected in parallel to both sides of the PD controller, the structural parameters of the IT2FNN start the self-learning process, and after the learning time of about 1.7 seconds, the output signal of the IT2FNN replaces the PD controller as the main output source of the control signal, which is the same as the result of Simulation 1. The above simulation results demonstrate the effect of the learning process of SLIT2FNN on the controller output and the improvement of the control accuracy in the second operating cycle with the parallel intelligent control structure.

It is important to note that the SLIT2FNN controller proposed in this chapter does not require any information about the mathematical model of the controlled object. In practice, there are a variety of intrinsic and extrinsic uncertainty effects, and if the PD controller used by the robot is tuned for one operating environment to achieve optimal control, it does not mean that the same control quality can be obtained in other operating environments. It is impractical
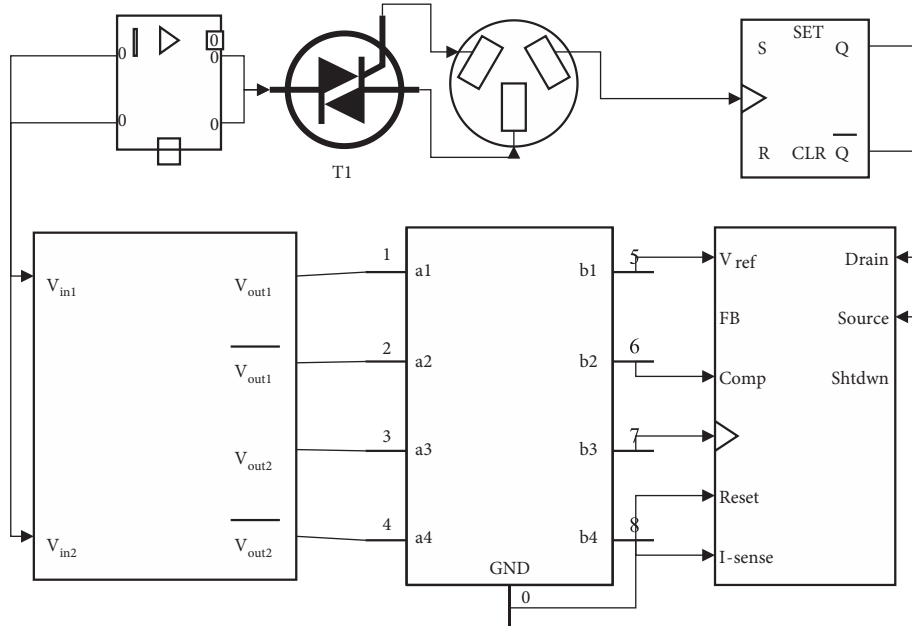
FIGURE 3: Fuzzy logic control system for Delta's robot.
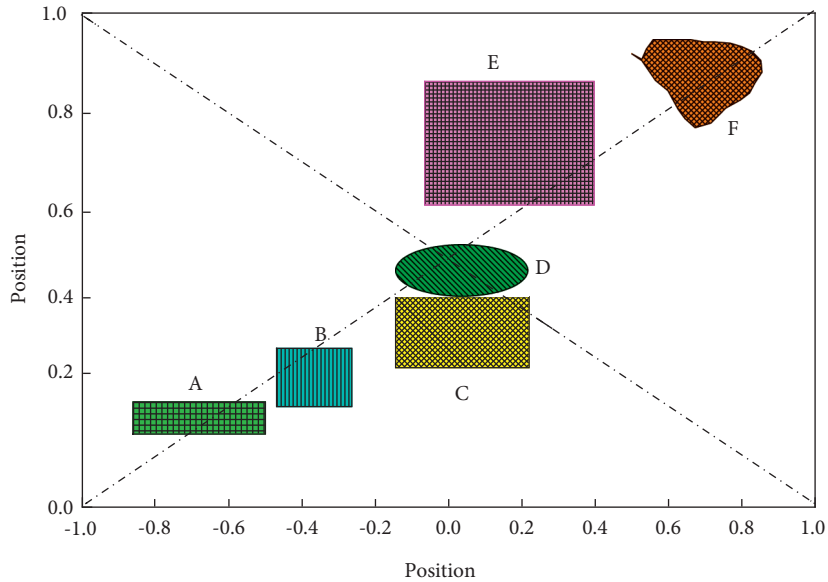
FIGURE 4: Trapezoidal affiliation function fuzzification method.

to perform optimal adjustment of PD control parameters for each case. Therefore, the parallel control system structure with self-learning capability proposed in this paper is of more practical value. This controller with parallel structure can not only provide the control effect of the corresponding PD controller when the neural network learns, but also make the control accuracy and robustness improvements to a great extent after the learning is completed and have higher environmental adaptability. At the same time, the parallel control structure proposed in this paper makes it easy to be applied to existing automation equipment and can greatly save hardware upgrade costs.

4.2. Neural Network Motion Results Analysis. To verify the optimized trajectory reuse algorithm, this section constructs the optimized trajectory database generated by the EFPF-PSO algorithm into a data index structure with two levels and six-leaf nodes using a vocabulary tree. Due to the limitation of the computer's computing power and time relationship, this algorithm only considers the case of 1 and 2 obstacles in the workspace, each database contains 180 sets of data, and the initial motion trajectory of the robot is set the same as before. This algorithm can generate corresponding optimized trajectory databases for specific application scenarios in practical applications and index the
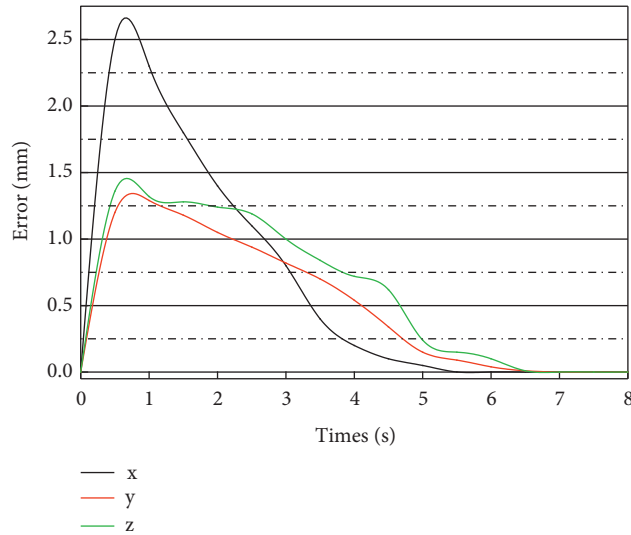
FIGURE 5: TCP position error using a controller.

databases according to the actual situation by classification. Firstly, the EFPFDB of the optimized trajectory is clustered for the first time by K-mean clustering algorithm; considering the characteristics of the number of obstacles, here the classification value $K = 2$. The clustering centers of the first feature layer of the database are C1 and C2, as shown in Figure 7. The algorithm has obtained the relationship between the location features of obstacles by unsupervised learning, and the EFPF values are classified.

The RPSOMP method consists of a PSO-based motion trajectory optimization algorithm and an optimized trajectory reuse method. The input is the trajectory to be optimized, and its EFPF vector is calculated first, and the optimized trajectory adjustment with the highest similarity EFPF stored in the database is calculated by vocabulary tree query. The preliminary optimized trajectory after the query is obtained using equation (9), and if the obtained trajectory is a feasible solution, the obtained trajectory is directly used for motion planning as the output result. This paper uses the center of sets type reducer to reduce the type of the ignition output set obtained. Since the antecedents and consequents used in this paper are both interval type-two fuzzy sets, the reduced-type set obtained is still an interval set. If the obtained optimized trajectory is not a feasible solution, the optimized trajectory given by the algorithm is used to initialize the PSO particles. The adjustment amount of the query result trajectory is used as the mean value of the particle initialization distribution, and the value calculated from the query result is used as the standard deviation to initialize the particle positions of the PSO and perform the optimization calculation. Finally, the adjustment amount of the obtained optimized trajectory and the EFPF of its initial trajectory are stored in the corresponding database, and the optimized trajectory is output. The method flowchart of the RPSOMP method is shown in Figure 8.

From the analysis in Figure 8, RPSOMP converges faster than EFPF-PSO, taking only 0.43 seconds on average, which fully satisfies the requirements for trajectory optimization calculation in engineering applications. Meanwhile, it should be emphasized that, in practical applications, if the data volume of the optimization database is large and the database is designed for a specific problem, in most cases, the RPSOMP method does not require repetitive optimization calculations but can directly use the retrieval results to generate optimization trajectories. This can further improve the time to obtain the optimized trajectory to about 0.001 seconds and meet the demand of real-time trajectory optimization. Although the EFPF-PSO algorithm has a chance to obtain optimized trajectories with smaller fitness values, not every calculation converges to the minimum fitness value, and its standard deviation is relatively large. In contrast, the results obtained using the RPSOMP method are more stable, have smaller standard deviations, and converge faster. For the robot online collision detection problem, a simplified characterization method of robot and obstacle in space is proposed. Based on this, the robot trajectory optimization problem is investigated using the PSO-based motion trajectory optimization algorithm. The demand trajectory of TCP is converted into the demand angle of driving joints and input into the controller through the inverse kinematics calculation. The input of the controller is the difference between the required angle of the drive joint and the actual angle and its derivative. As the IT2FLC design benchmark, T1FLC uses a type 1 fuzzy set of trapezoidal membership function as the antecedent and a type 1 fuzzy set of triangular membership function as the subsequent piece.
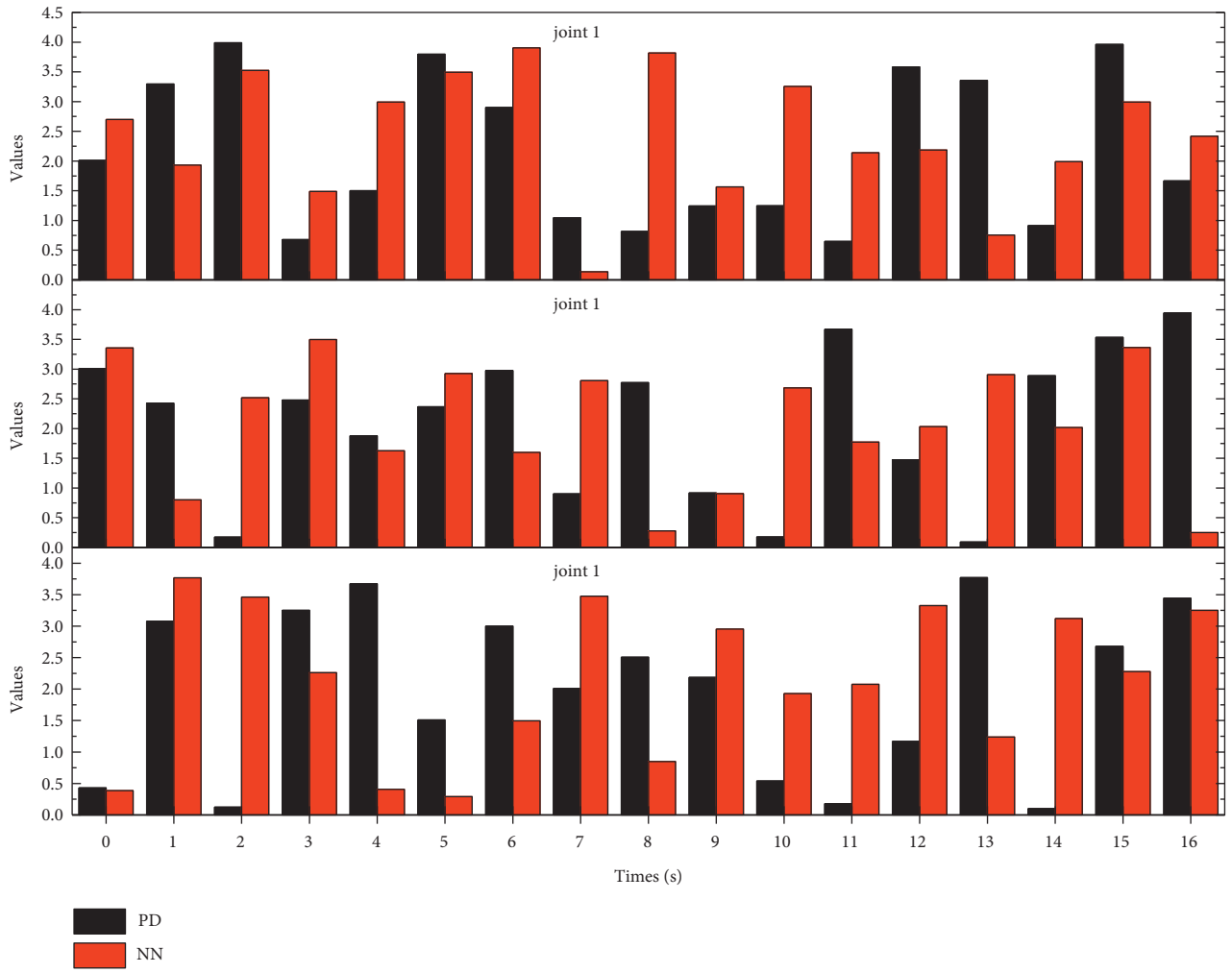
Figure 6: Control signals.
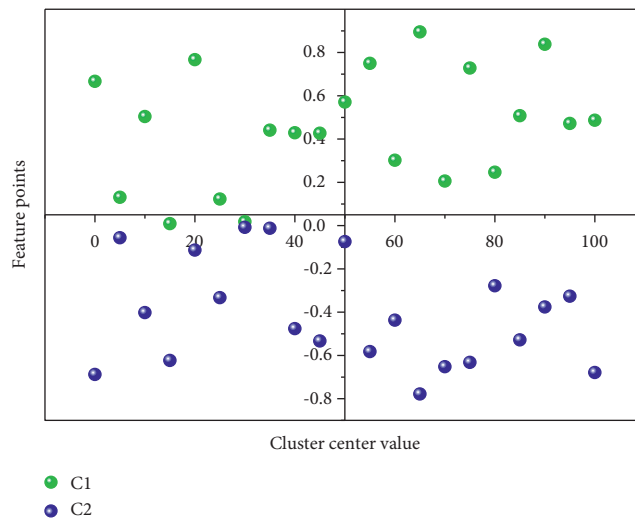
Figure 7: Clustering center of the first feature layer of EFPFDB.
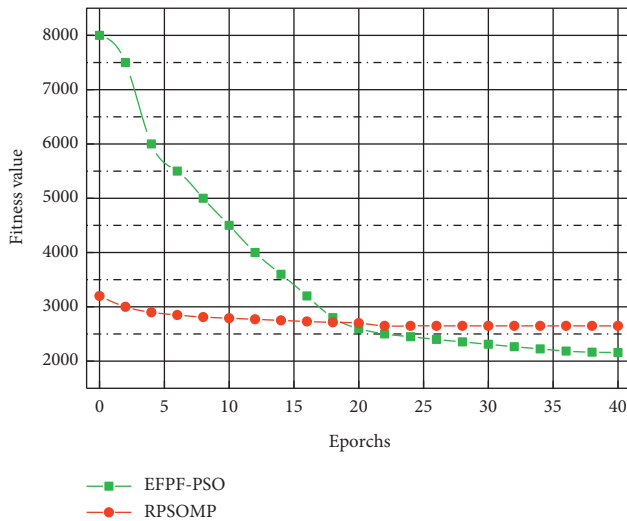
Figure 8: Average of the adaptation of RPSOMP and EFPF-PSO.

5. Conclusion

This method is based on the optimized T1FLC, and the fuzzy degree is used to fuzzify the type-one fuzzy subordinate function into the interval type-two fuzzy subordinate function to obtain the IT2FLC. Three fuzzification methods for the trapezoidal subordinate function are studied, and the relationship between the fuzzy degree, the fuzzification method, and the output control surface is given. OSEC is used to ensure the control accuracy of IT2FLC after fuzzification, and the calculation and value of OSEC are given. A SLIT2FNN with a seven-layer structure is established, and a parallel intelligent controller based on SLIT2FNN is proposed on this basis. The transient performance during the learning process of the neural network can be compensated by the parallel PD controller, which makes it easy to upgrade the conventional controller. In the design of the front piece set of SLIT2FNN, the dual sequence symmetric trapezoidal affiliation function arrangement method is used to make the system's self-learning law and stability analysis in analytic form. In the design of the neural network self-learning law, a parameter self-learning algorithm based on the SMC theory is established to adjust the structural parameters of IT2FNN online. An industrial intelligent robot software system based on the robot operating system architecture is established, and the kinematic and dynamic models, mechanism parameter optimization algorithms, SLIT2FNN-based intelligent controller, and RPSOMP method are implemented based on this paper. Based on the research in this paper, the intelligent task planning module is further investigated to make the robot autonomous and capable of operating in unstructured spaces and environments and completing nondeterministic tasks online in real time.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] K. Vonsevych, M. F. Goethel, J. Mrozowski, J. Awrejcewicz, and M. Bezuglyi, "Fingers movements control system based on artificial neural network model," *Radioelectronics and Communications Systems*, vol. 62, no. 1, pp. 23–33, 2019.

[2] P. Subbash and K. T. Chong, "Adaptive network fuzzy inference system based navigation controller for mobile robot," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 2, pp. 141–151, 2019.

[3] W. He and Y. Dong, "Adaptive fuzzy neural network control for a constrained robot using impedance learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1174–1186, 2017.

[4] J. H. Jeong, K. H. Shim, D. J. Kim, and S. W. Lee, "Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 5, pp. 1226–1238, 2020.

[5] F. Jiang, F. Pourpanah, and Q. Hao, "Design, implementation, and evaluation of a neural-network-based quadcopter uav system," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 3, pp. 2076–2085, 2019.

[6] J. Fei and T. Wang, "Adaptive fuzzy-neural-network based on RBFNN control for active power filter," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 5, pp. 1139–1150, 2019.

[7] M. Fouzia, N. Khenfer, and N. E. Boukezzoula, "Robust adaptive tracking control of manipulator arms with fuzzy neural networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 6131–6141, 2020.

[8] J. Narayan, E. Singla, S. Soni, and A. Singla, "Adaptive neuro-fuzzy inference system-based path planning of 5-degrees-of-freedom spatial manipulator for medical applications," *Proceedings of the Institution of Mechanical Engineers—Part H: Journal of Engineering in Medicine*, vol. 232, no. 7, pp. 726–732, 2018.

[9] A. M. Rao, K. Ramji, B. S. K. Sundara Siva Rao, V. Vasu, and C. Puneeth, "Navigation of non-holonomic mobile robot using neuro-fuzzy logic with integrated safe boundary algorithm," *International Journal of Automation and Computing*, vol. 14, no. 3, pp. 285–294, 2017.

[10] M. Mirzaey, M. B. Jamshidi, and Y. Hojatpour, "Applications of artificial neural networks in information system of management accounting," *International Journal of Mechatronics, Electrical and Computer Technology*, vol. 7, no. 25, pp. 3523–3530, 2017.

[11] F. Farivar, M. S. Haghighi, A. Jolfaei, and J. Alazab, "Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial IoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2716–2725, 2019.

[12] J. Li, Z. Li, Y. Feng, Y. Liu, and G. Shi, "Development of a human-robot hybrid intelligent system based on brain tele-operation and deep learning SLAM," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 4, pp. 1664–1674, 2019.

[13] C. Sun, H. Gao, W. He, and Y. Yu, "Fuzzy neural network control of a flexible robotic manipulator using assumed mode method," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5214–5227, 2018.

[14] P. M. Kebria, A. Khosravi, S. Nahavandi, D. Wu, and F. Bello, "Adaptive type-2 fuzzy neural-network control for tele-operation systems with delay and uncertainties," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 10, pp. 2543–2554, 2019.

[15] I. Prasojo, P. T. Nguyen, and N. Shahu, "Design of ultrasonic sensor and ultraviolet sensor implemented on a fire fighter robot using AT89S52," *Journal of Robotics and Control (JRC)*, vol. 1, no. 2, pp. 55–58, 2020.

[16] M. M. Ferdaus, S. G. Anavatti, M. Pratama, and M. A. Garratt, "Towards the use of fuzzy logic systems in rotary wing un-manned aerial vehicle: a review," *Artificial Intelligence Review*, vol. 53, no. 1, pp. 257–290, 2020.

[17] J. P. C. de Souza, A. L. M. Marcato, E. P. de Aguiar, M. A. Jucá, and A. M. Teixeira, "Autonomous landing of UAV based on artificial neural network supervised by fuzzy logic," *Journal of Control, Automation and Electrical Systems*, vol. 30, no. 4, pp. 522–531, 2019.

[18] C. R. Thilahar and R. Sivaramakrishnan, "Fuzzy neuro-genetic approach for feature selection and image classification in augmented reality systems," *IAES International Journal of Robotics and Automation*, vol. 8, no. 3, pp. 194–204, 2019.

[19] Z. Li, R. Xu, P. Cui, and S. Zhu, "Artificial neural network based mission planning mechanism for spacecraft," *International Journal of Aeronautical and Space Sciences*, vol. 19, no. 1, pp. 111–119, 2018.

[20] D. H. Dos Reis, D. Welfer, M. A. De Souza Leite Cuadros, and D. F. T. Gamarra, "Mobile robot navigation using an object recognition software with RGBD images and the YOLO algorithm," *Applied Artificial Intelligence*, vol. 33, no. 14, pp. 1290–1305, 2019.

Research Article

Smart City + IoT Standardization Application Practice Model and Realization of Key Technologies

Chunqiao Song [iD] [1] **and Xutong Wu** [iD] [1,2]

[1]*Faculty of Innovation and Design, City University of Macau, Macau 999078, China*
[2]*Macau University of Science and Technology, Institute of Social and Cultural Research, Macau 999078, China*

Correspondence should be addressed to Xutong Wu; u17092105073@cityu.mo

Smart city gathers heterogeneous information, which requires a unified data access and collection mechanism. In this paper, based on the computing framework of IoT put forward to satisfy urban sensing network data acquisition, transmission, and computing technology system, according to the function requirements of IoT gateway, we put forward the method of building software system and hardware system and the urban perception network middleware architecture that can meet the demand of M2M, which can realize dynamic management of cognitive resources. It also supports ubiquitous services over HTTP. This is of great significance for providing intelligent services with diversity of solvers, controllers, and computing terminals. This paper focuses on how to comprehensively apply intelligent sensing, wireless transmission, data mining, and other technologies of the Internet of Things to associate data with typical application examples such as vehicle sensing and positioning technology and urban security sensor application. These typical applications can effectively solve the practical problems in urban operation and management, meet the needs of the expansion of the concept of Internet of things in urban management and the continuous deepening of management, and improve the scientific and intelligent level of urban operation and management.

1. Introduction

As the trading center and gathering center of human beings, city is the product. The emergence of cities is the symbol of human society entering the civilized era, and also the advanced form of human group life. The continuous acceleration of urbanization has led to the rapid increase of urban population, the rapid expansion of urban scale, and the continuous enhancement [1]. It improves the level of urban services and promotes the sustainable and leapfrog development of cities through comprehensive and transparent information perception, wide and safe information transmission, and intelligent and efficient information processing. In this way, a new form of urban development can be constructed, so that the city can automatically perceive and effectively make decisions and achieve control [2]. As the infrastructure of smart city, information technology is undoubtedly composed of the three most core technology hot spots of the Internet of Things, cloud computing, and data

association technology. In fact, these three technologies all belong to the general description of the platform including many technical branches. Internet of Things technology [3] is based on Radio Frequency Identification RFID and other sensing devices.

From the perspective of information technology, "wisdom city" intelligent information technology in the city has "four" characteristics: ubiquity, utility, intelligence, and greening [4]. (1) Ubiquity: the application of information technology is everywhere, people and nature, man and society, and nature and society are linked in the pan in the network information for smooth circulation and fusion. The network has become the basic production tool of the information society [5]. (2) Utility: the deployment, application, and management of IT infrastructure will be the same as those of water, electricity, and gas, with the principle of "centralized service, on-demand use, virtual ownership and easy to use" [6]. (3) Intelligence of infrastructure and its applications is more convenient. Information sharing

technology under data fusion and application system with intelligent perception and respect for user experience affect all aspects of people's lives. Smart city will shape the vision of a better life "benefiting everyone" [7]. (4) Flexibility: "ubiquitous information acquisition equipment; unimpeded information transmission channels; information sharing under security guarantee; powerful data processing centers; and intelligent software and services play more roles in urban operation than hardware production and manufacturing [8]. Concepts such as "smart Earth" and "smart city" have emerged. "Smart city" is a new urban form with learning, adaptation, and innovation capabilities. Based on the network integration of Internet of Things, telecommunication network, wireless network, and other networks, and using a variety of high-tech tools, it establishes an urban ecosystem including government, citizens, and commercial organizations [9]. It can be seen that the Internet of Things actually includes three entities: human, machine, and object. Intelligent devices are used to sense and acquire information about objects and their surroundings through recognition and perception technology [10], and then information and knowledge are transmitted through network transmission and communication technology and equipment. Finally, through the intelligent message retention, and knowledge processing technology and equipment to achieve intelligent management and control of objects or the physical world of a kind of "object perception, people interconnection, everyone interoperability" efficient, intelligent network, database and intelligent computing technology (including artificial neural network and evidence combination), standardization technology, etc. [11], literature [12] applies sensor data fusion technology to activity recognition in the field of smart home. Considering the advantages of intelligent computing, researchers gradually design and use intelligent algorithms to carry out data fusion research. For example, the improved particle swarm optimization algorithm is applied to fuse the multisensor information in the Internet of Things for precise measurement [13]. Based on the event-driven central point fusion algorithm, a fusion timing control mechanism based on differentiated services is proposed [14]. Cloud computing, comprehensive perception, intelligent computing analysis, ubiquitous connectivity, digitalization, information, and network of modern cities are new in recent years due to the integration and convergence [15]. The concept of "smart Earth" proposed by IBM elevated smart city construction into a national strategy [16]. Next, with wisdom city construction mainly aimed at people widespread concern in the construction of the city of urban traffic congestion, public security, network optimization deployment and nondestructive testing, digital city life, medical care, urban greening, and sudden disasters, etc. [17] make full use of the Internet of things technology in the intelligent sensing and comprehensive perception, information sharing, and service advantages, promote the improvement of modern social civilization and government management and service level, and comprehensively improve the happiness index of urban and rural residents [18, 19]. Through the construction of urban operation monitoring platform based on the Internet

of things, realize the combination of automatic and manual new regulatory model, effectively improve the government guarantee public security and the ability to deal with emergencies, maximum limit to prevent or reduce damage [20, 21], safeguard the life property safety of the public, safeguarding national security and social stability, and promote social economic comprehensive, harmonious, and sustainable development. The combination of the Internet of things and urban management is getting closer and closer. At present, there are application solutions of the Internet of things technology in urban management, environmental monitoring, logistics informatization, intelligent transportation, public security, and many other fields [22, 23], and demonstration work has been actively carried out in all aspects of the city [24].

This article first analyzes the metropolitan area network perception function definition and hierarchy, and then according to the data acquisition and exchange demand of the city of wisdom, this paper proposes a IoT gateway hardware and software of the system, then considers wisdom city service framework based on cloud computing, and discusses the design principle of metropolitan area perception network middleware and implementation method of the system architecture. Finally, the above main architecture is verified by performance test experiments. Using the theory of innovation and technology to build a series of specific function models, this part will analyze the diverse needs of construction of urban management wisdom, using the IFC standard, data mining, and text mining method, combined with a variety of emerging information technologies to achieve the extraction of submodels, definition, and integration, such as the generation process; this is part of the model to study further on. It is also a key link between the past and the future.

2. Key Technology System Architecture of Smart City Internet of Things Application Practice Model

2.1. Application Practice Model of Intelligent City Internet of Things Research Technical Framework. The technical architecture of a typical Internet of Things system includes three parts: perception layer, transmission layer, and application layer. This training system adopts typical architecture design, and the functions of each layer are as follows:

2.1.1. Perception Layer. Look for the training function, cultivating students perception layer function, the characteristics of a variety of sensors, detection methods, cognitive, and selection ability, for the application of single chip microcomputer (51 series and CC25xx series) programming ability, and the business requirements and the engineering application environment, according to the system planning of the node deployment method and the construction deployment according to specification.

2.1.2. Transport Layer. From the perspective of system composition and technical functions, and from the perspective of practical training function, the transmission layer

focuses on cultivating students' skills in the configuration and application of Internet of Things products, as well as the application development ability of SCM (CC25xx series) or embedded processor.

2.1.3. Application Layer. The smart City + IoT traffic scheduling system is used to monitor commuting at important traffic crossings (such as Bridges) and control the operating status of traffic lights at corresponding locations. The system composition is shown in Figure 1. Subsystem operation and lighting control subsystem, the user through a Web application software, set up bridge state of overload, the linkage of the road traffic status, traffic lights to run the action control parameters, collection and service software issued instructions according to requirements of the linkage control and feedback data, through a Web application software system running state.

The concept of smart city IoT standardization is based on the premise of urban IoT standardization informatization and the development of emerging information technology in China. It is an effective means to realize the whole life cycle management of construction projects, which can promote the optimization of urban IoT standardization related professionals and personnel and change the content and mode of construction project management. Innovation and integration of a variety of emerging information technology are the core of wisdom city of standardization of the Internet of things, but the wisdom city IoT standardization idea advocated is not only a collection of a large number of systems integration and information infrastructure, but also the whole life cycle of construction project management and multiple parties together, creating wisdom, integrated with the construction of human environment, providing convenient, efficient, and efficient project management methods. In order to achieve these goals, the concept of smart city IoT standardization emphasizes the establishment of anthropomorphic sensing, connectivity, and intelligence functions, as well as harmonious coexistence with the construction environment to achieve sustainable development of construction projects. Based on the above analysis, the conceptual model of IoT standardization concept of smart city is obtained, and the causes, management objectives, main features, and realization mechanism of IOT standardization concept of smart city are revealed.

3. Internet of Things Contributes to the Refined Model and Key Technologies of Urban Management

Effective supervision of movable objects (breakfast carts, newsstands, etc.) and frequently changed objects (advertisements, plaques, etc., of shops on both sides of the street) can make full use of radio Frequency Identification (RFID) tag positioning technology, the core technology of the Internet of Things. The legal and compliant supervision objects shall be installed with identifiable electronic tags. The vehicle-mounted monitoring and law enforcement system is established based on the integrated supervision functions of

base stations, panoramic cameras, and other equipment to realize vehicle-mounted patrol, automatic perception, real-time positioning, and rapid investigation and evidence collection of supervised objects and improve the efficiency and intelligence level of management. The supervision mode is shown in Figure 2.

With the continuous improvement of people's living standard, city parks, temple fairs, or the famous tourist attractions such as area in a particular time period (such as holidays or during certain activities frequently, people "blowout" phenomenon, it also makes the crowds gathered risk further ascension and increased security hidden danger, and the crowd safety management difficulty also will increase. Traditional management mainly adopts the way of on-site monitoring with a large number of manpower, but it can only play the role of alarm disposal and lacks the prediction and early warning of crowd gathering risk.

$$f_p = \min \sum_d \frac{D_J}{P}. \tag{1}$$

(1) In the multiobjective optimization model, both models have shortcomings in measuring progress in improving optimization. Multiobjective Selection Based on Dominated Hypervolume (SCPIMM) model showed better performance and optimized results for different front shapes.

(2) According to the group based chain coding method, the parent individuals were initialized to generate a parent population of size N, and the fitness of individuals in the population was evaluated.

(3) A rapid nondominant ranking was carried out on the evaluated parent population to classify the dominant rank of individuals in the population.

(4) Population evolution operation: the algorithm adopts the tournament selection method of two parent individuals selected from the population, and to improve the two parent individuals type single point cross to generate a hybrid individuals, and then to intelligent hybrid individual variations, get a child individual, pair the last individual fitness evaluation, and add the individual child population pool. Repeat this process until the subpopulation size is N.

(5) Optimal N individuals were selected as the next generation parent population by using hypervolume-based population renewal mechanism for parent population and child population.

4. Information Model Construction of Intelligent City Internet of Things

The modeling process should reflect the main features of the standardization concept of the Intelligent city Internet of Things, integrate a variety of emerging information technologies into the model, and meet the needs of modern construction project information management. SCPIMM can not only meet the individual needs of different participants in the process of construction project management
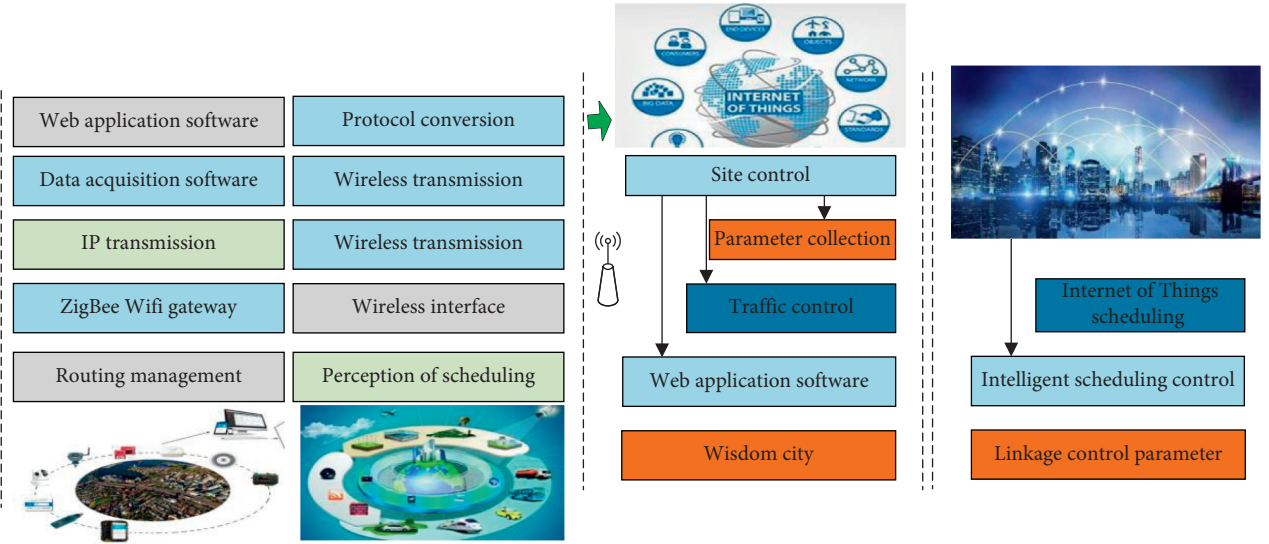
Figure 1: Block diagram of smart city + Internet of Things traffic scheduling control system.
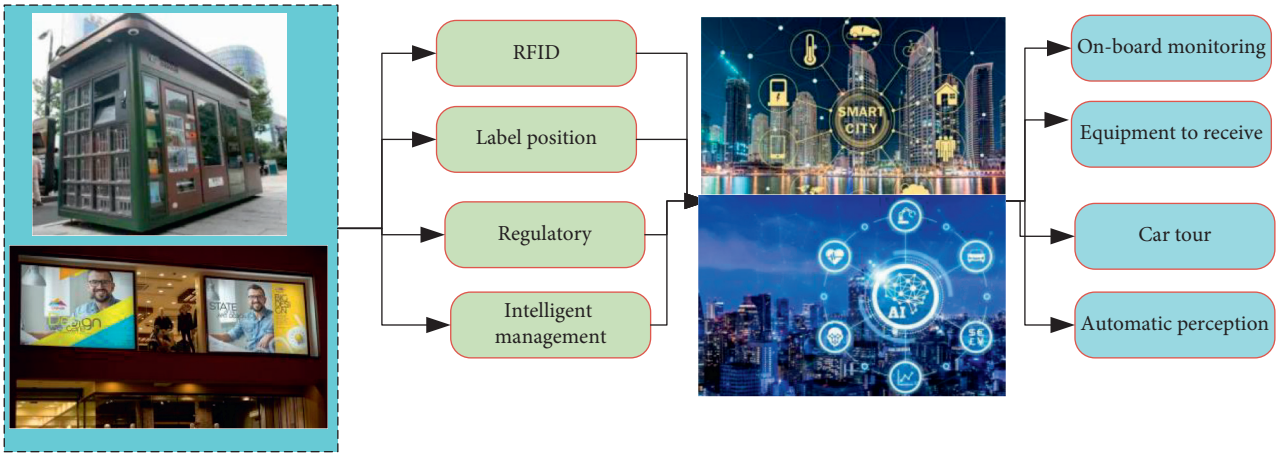
Figure 2: Schematic diagram of vehicle-mounted monitoring law enforcement methods.

for information management, but also realize the effective management and control of the whole life cycle information and improve the efficiency and level of construction project information management. The basic framework structure of SCPIMM proposed in this paper is shown in Figure 3.

(1) Data layer can be divided into structured, unstructured, and process information according to the structural characteristics of construction project information. For different information structures, SCPIMM has established three kinds of databases to realize classified storage and management: the storage and management of structured information are carried out through the structured information database constructed by the model. The unstructured information is stored and managed through the unstructured information database constructed by the model by establishing the link relationship between the model and the document. The process information of construction project is stored and managed by process information database according

to the principle of mapping. According to the characteristics and practical experience of SCPIMM, the model database serves as the carrier and storage terminal of project information and takes information exchange interface and middleware technology as the key to process and process heterogeneous information, providing support for other levels of the model.

(2) In the process of construction project information management, the platform layer uses sensors and information perception technology to build an information management platform with perceptive ability and real-time information interaction, so as to realize the collection and monitoring of construction project information resources and complete the information management work of the whole life cycle of the model. The platform layer of the model is composed of a variety of sensors, generally including HD camera, concentration sensor, temperature and humidity sensor, two-dimensional code label, radio
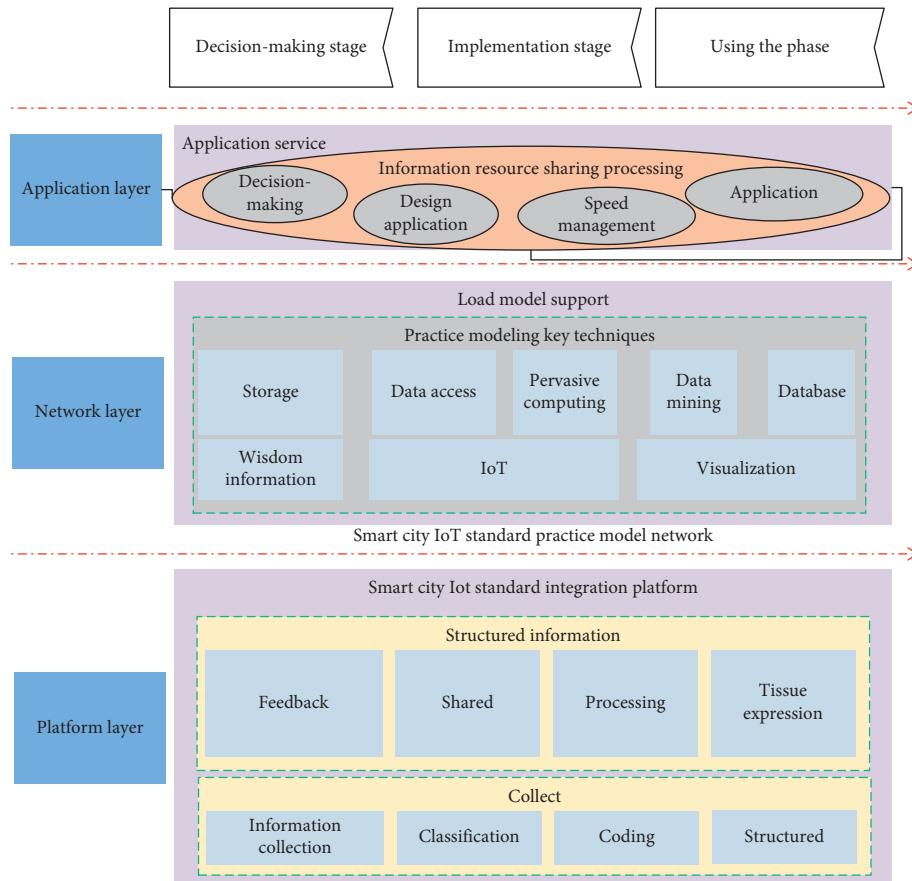
FIGURE 3: Main frame of IoT standardization information model for smart city.

frequency label reader, GPS, and other sensing terminal devices. According to different stages and applications of the whole life cycle of construction, the platform layer can classify and encode all the collected information and process the information according to structured, unstructured, and process information classification.

(3) The network layer of the network layer model can obtain and process the information provided by the platform layer and build a network system by utilizing high-performance computing, data mining, visual management, and other functions provided by emerging information technologies such as the Internet of Things, ubiquitous computing, and 4D visualization. Its core capability is the transmission and processing of information. Information network construction, based on the technology of Internet of things, can be achieved from a single project to the whole construction project information management of specific equipment, the whole process of information interaction, from the construction project detailed information to the terminal transmission path through the network, sensor, information storage database, to the network infrastructure, realizing information sharing and interaction.

(4) The application layer of the application layer model uses a variety of emerging information technologies to form a whole-process intelligent information management process from event prediction to decision assistance and emergency linkage and to create a comprehensive management system of information collection, classification and coding, organization and expression, processing and processing and transmission, sharing, and feedback. The application layer of the model is composed of application software with different functions, including investment analysis, architectural design, structural design, cost management, schedule control, quality control, safety management, and energy consumption analysis. The application layer of the model is the interface between the network layer and the user, which combines with the user's needs and adopts P2P access mechanism based on the network topology structure to provide different services for different participants.

The design of information extraction mechanism is the premise of submodel construction, which is directly related to the efficiency of information management of construction projects. The main purpose of information extraction is to obtain the information sample similar to the information needed at present. In this paper, the submodel information

extraction mechanism is designed according to the similarity principle. The following three methods are adopted to define the similarity between the target and the sample in the extraction process.

(1) Similarity between information is defined by Euclidean distance P.

$$d_p = \sum_i W_i (p_{\text{ok}} - p_i)^2. \tag{2}$$

(2) Similarity of information is defined by Manhattan distance.

$$d_p = \sum_P W_p |p_{\text{ik}} - p_i|. \tag{3}$$

(3) Similarity between information is defined by infinite modulus distance.

$$d_p = \max W_i |p_{\text{ik}} - p_i|. \tag{4}$$

In the information preparation stage of the submodel, massive construction project information should be preprocessed by stem extraction (Stemming, stem segmentation, stop words removed, punctuation deleted, etc.). For English, stem should be extracted, and nouns, adjectives, conjunctions, and adverbs should be taken as feature items. For Chinese stem segmentation, namely, word segmentation, specific word segmentation techniques include maximum matching, reverse maximum matching, and traversal one by one. The feature items after word segmentation should be processed by dimensionality reduction, that is, selecting the required feature items f to form a set N.

$$p(d, t) = \frac{\lg(d, f)(N/n_j)}{\sqrt{\sum (d, t)^2 \lg(N/n)^2}}. \tag{5}$$

The application layer of associated data is mainly composed of Resource Description Framework (RDF) triplet evaluation, data visualization, semantic query and browsing, personalized information push, and other modules. First, the associated data automatically generated by the system is input to the application layer from the data storage and index layer. In RDF triplet evaluation module, experts and users score automatically generated relational data. The associated data retained after evaluation is displayed in the data visualization module through API isoform. For semantic query, browsing and other requirements in the application, semantic query, and Tabulator plug-in in browsing module can be realized. In addition, the personalized information push module is responsible for pushing the content that users may be interested in according to customized information.

5. Results and Analysis

The historical monitoring data accumulated by thousands of sensors in cities during the civil period can be an important part of the big data of flood control monitoring. On the basis

of full investigation and investigation, special analysis of monitoring data and related data of other cities in the same period was carried out to find the consistency, concurrency, and derivative relationships among them, find the spatial distribution rules, and put forward scientific and reliable suggestions to better support flood control and emergency response work. For example, based on rainfall monitoring data, corresponding flood control risk assessment and early warning models can be established, and data mining and other means can be used to further improve the ability to master the law between rainfall and urban management cases, so as to realize the refinement of daily urban management related to flood control. Figure 4 shows the analysis results of significant correlation between rainfall monitoring data and some types of management cases.

In terms of enabling the number of physical nodes, the number of physical nodes occupied by SCPIMM-based VM (Virtual Machine) deployment algorithm is almost the same as that occupied by CGA-based VM deployment algorithm and is smaller than that occupied by the priority matching heuristic (PMH) deployment algorithm, physical nodes. The smaller the number of tablets, the more the energy saving. In terms of the overall cluster load performance, the overall cluster load variances of the SCPIMM-based VM deployment algorithm are smaller than those of the two VM deployment algorithms. The smaller the load variances are, the better the load balancing effect of the server cluster is. Based on the above analysis; see Figures 5 and 6.

To compare the network performance of the Structure-Conduct-Performance Interacting Multiple Model (SCPIMM) protocol with that of Low Energy Adaptive Clustering Hierarchy (LEACH) and SCPIMM, we first compare the clustering formed by the three protocols, as shown in Figure 7. It can be seen from the figure that the number of clusters formed by LEACH is too large, which is caused by the randomness of the selection of cluster heads. In LEACH, each sensor node decides whether to become a cluster head according to a certain probability, which has a strong randomness. The clustering formed by SCPIMM is better than that formed by LEACH, but it is different in size compared with SCPIMM. This is because the SCPIMM give full consideration to the state of the network to determine the optimal number of cluster heads, and then according to the residual energy of nodes, the distance between nodes and the distance between nodes and base stations to determine the optimal selection of cluster heads, thus forming clumps and uniform distribution, keep the optimal number of cluster heads, reduce the energy consumption of network, and improve the performance of the system.

The data received by the base station (sink node) is shown in Figure 8. We can see from the table that the proposed SCPIMM can receive more data, because we consider the residual energy of nodes when clustering, within the cluster distance and the distance with the base station, effectively reducing the energy consumption and reducing the amount of data transmission delay and improving the efficiency of data transmission, to send more data to base station. However, LEACH does not consider factors such as the remaining energy of nodes when
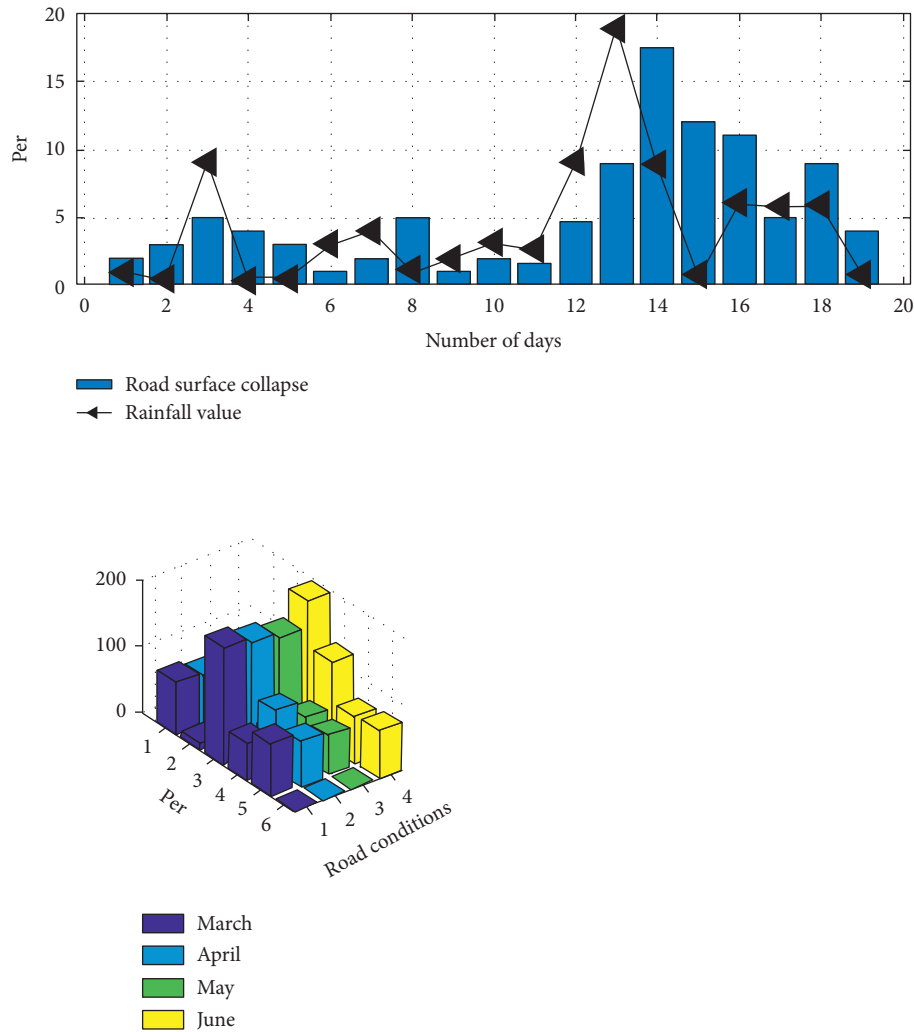
Figure 4: Correlation analysis of intelligent city IoT rainfall monitoring data and urban management case data.

clustering and randomly selects nodes as cluster heads. It is possible to select nodes whose energy is about to run out as cluster heads, resulting in data loss. Although SCPIMM considers residual energy and other factors when clustering, the optimal number of cluster heads does not change with the network state, and the influence of the distance between cluster heads and base station is not taken into account.

Based on the overall construction mode of smart city informatization service system, integrated Internet of Things construction mode, cloud computing construction mode, and intelligent infrastructure construction mode, and combined with the application demand of smart city people in smart city construction, the construction mode of smart city is explored and analyzed.

According to the effective data collected from the questionnaire, overall, smart city citizens have higher demands for smart government, smart environment, smart medical treatment, smart transportation, and smart family in smart city construction, while their demands for smart energy and smart logistics are relatively low. See Figure 9.

At the same time, from the point of view of intelligence security data, security is also a matter of great concern to the public, from the point of view of data, and the average score of security demand is 3.36 points. Among them, citizens' demands for early warning of natural disasters and emergency response of major emergencies account for a large proportion, accounting for 21% and 20%, respectively, as shown in Figure 10. The reason may be that the disasters such as earthquake, tsunami, and fire that have occurred in recent years are more serious, which make citizens have a kind of panic in psychology. They hope to take corresponding measures in the early warning of natural disasters and emergency response in case of major emergencies to avoid heavy costs. However, they did not pay much attention to the information tracking of registered population and permanent resident population, indicating that the citizens of smart cities can accommodate floating population, and the informatization of population and family planning in smart cities is also relatively good. Covering the food and drug traceability system, citizens are more concerned with the quality of the food and drug problems frequently recently, and also it is also a concern that citizens using the sensor monitoring food and drug trace the source of wisdom city that can be in shopping malls and supermarkets to
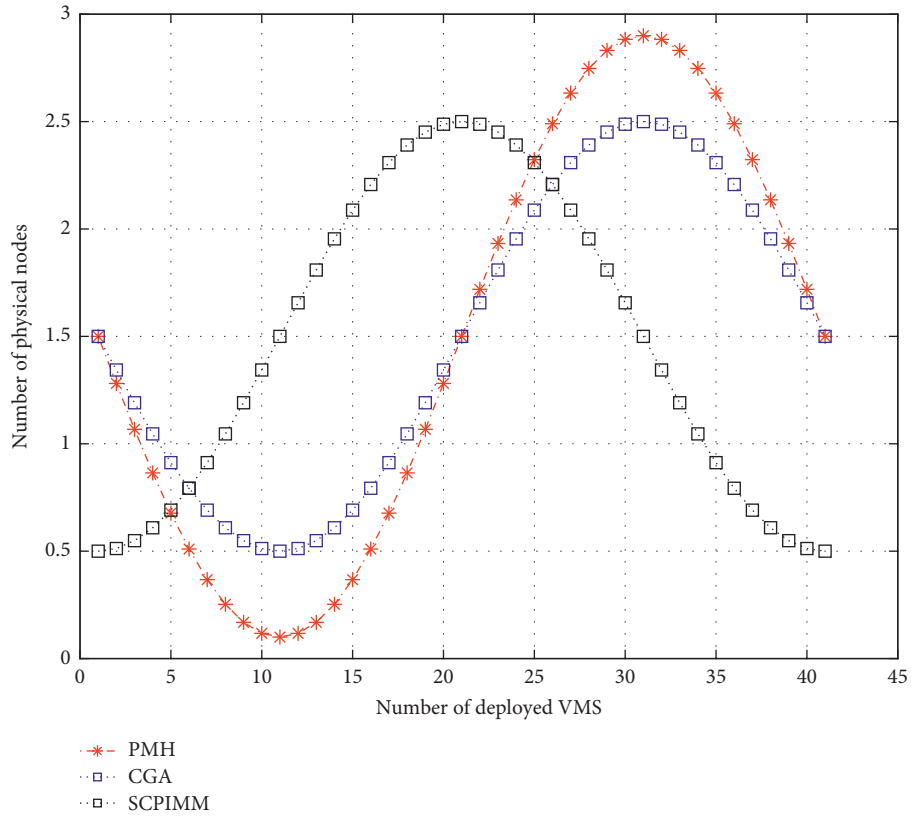
Figure 5: Schematic diagram of enabling physical node number.
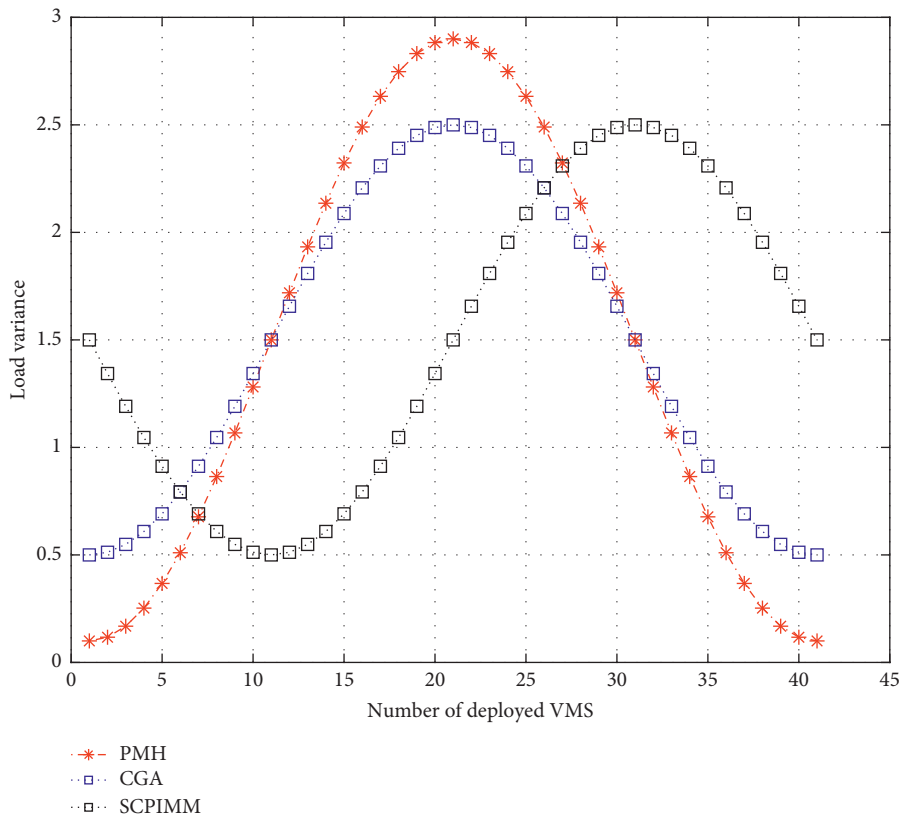
Figure 6: Schematic diagram of cluster load variance.

(a)

(b)
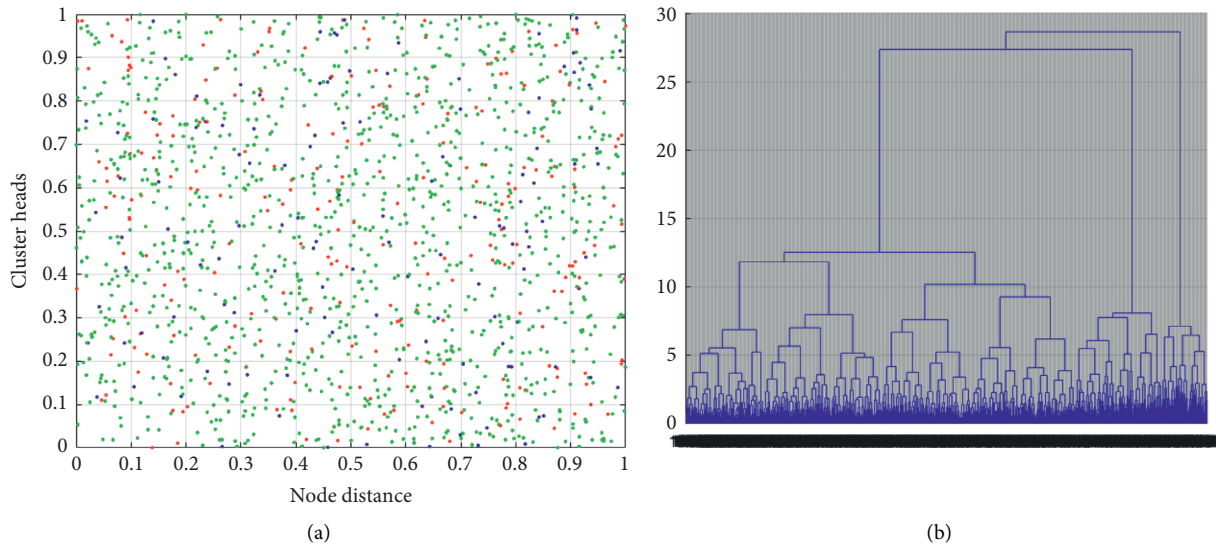
FIGURE 7: (a) Clustering formed by LEACH. (b) Clustering tree formed by SCPIMM.
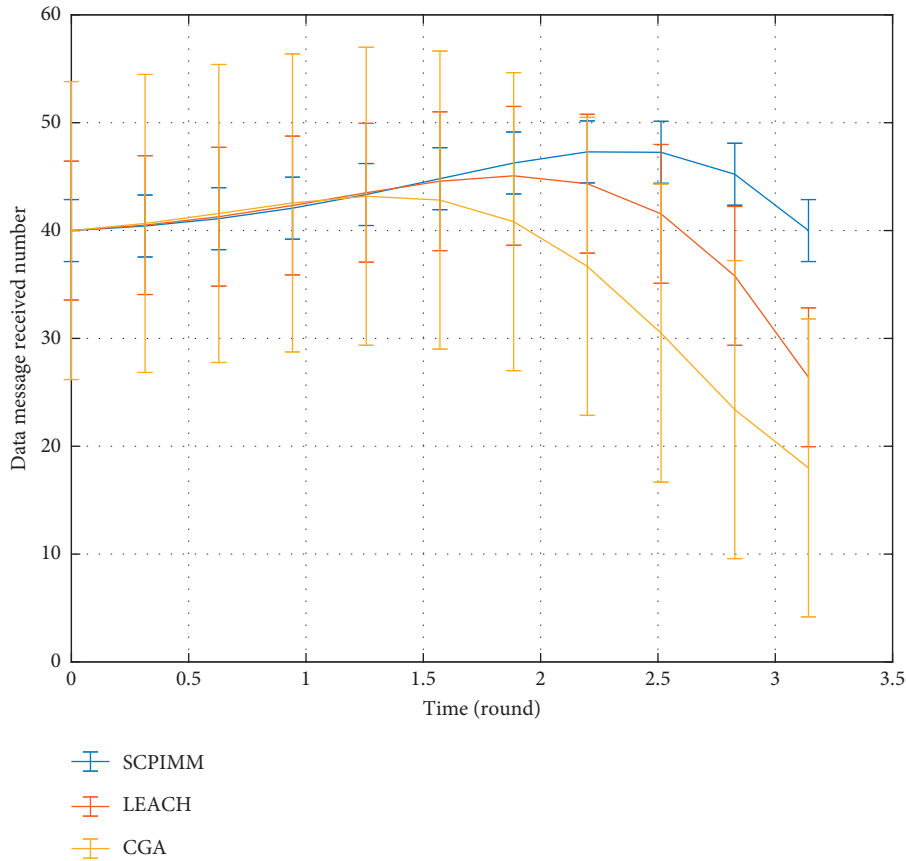
FIGURE 8: Data received by the base station.

install HD probe for video surveillance, clearly photographing food stand in front of the person's face, to prevent the poisoning of food and drugs.

The physical equipment layer is the infrastructure of a smart city, including access control devices, attendance devices, consumer machines, parking gates, and wireless sensor networks for sensing environmental changes. All devices access the information system of the data center through the Internet of Things gateway and communicate with the device abstraction layer of the intelligent building information system middleware. Middleware includes common service layer, sensor network management layer
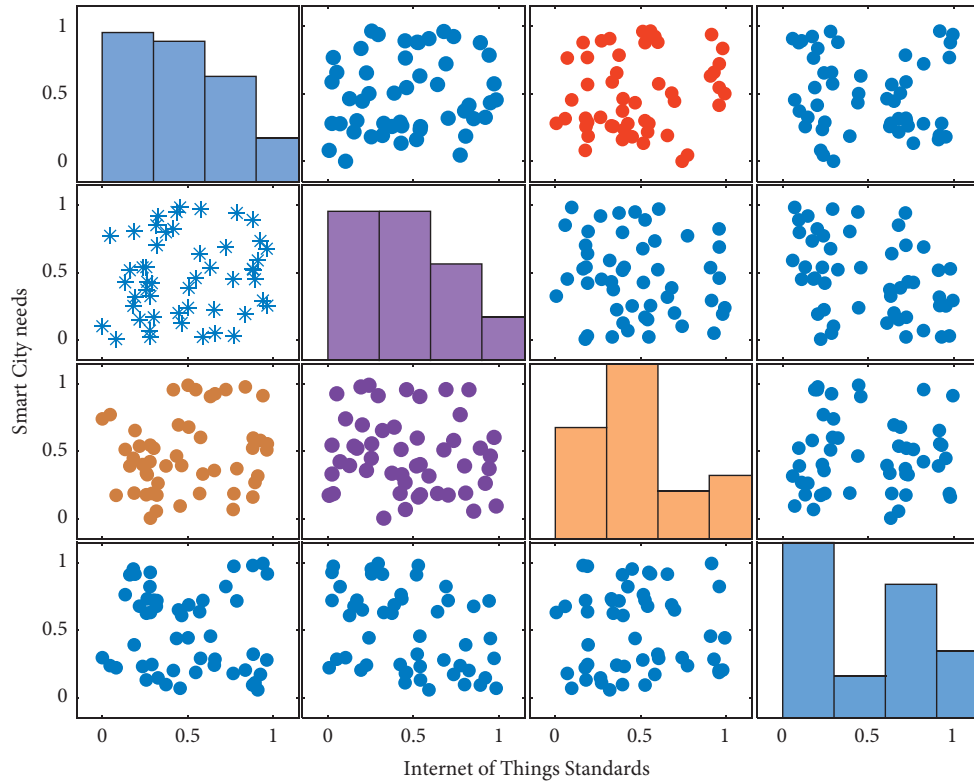
FIGURE 9: Application requirements of smart city construction.
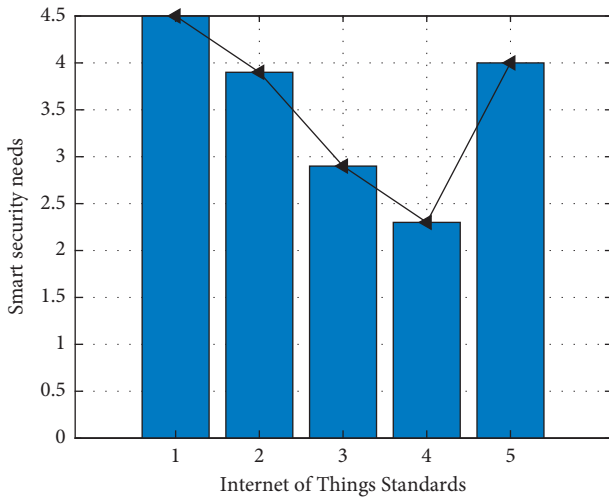
FIGURE 10: Intelligent security function requirements.

and application service layer. The sensor NMS topology consists of configuration services and abstraction layers. Topology and configuration services mainly realize the correspondence between physical space location of physical devices and information space location, and the interface of abstraction layer realizes the standardization of physical device access. The common service layer contains logging services, database services, and security services. The application service layer provides HTTP-based Web interfaces and local API interfaces for various applications.

6. Conclusion

At present, the urban management mode has undergone the transformation from digitalization and grid to urban operation, and its management direction is facing the development from digitalization to intelligence. These higher requirements for urban management and service work make the urban management must be based on the requirement of the people's fundamental starting point, adhering to the people-oriented, service for the people, to keep pace, through technology and management innovation to promote urban management work to a new level, to facilitate better break through the limitations of time and place, so that the people feel the role of urban management, enjoying the achievements of urban management. As an emerging technology integration means, Internet of Things technology can be applied in monitoring, transmission, analysis, and application of urban operation management information due to its characteristics of "intelligent identification, positioning, tracking, monitoring, and management" and has gradually become a bridge between "digital city" and "smart city." It will be the ultimate goal and strategic direction of urban informatization to establish an intelligent and refined management mode of the city. As a breakthrough of the bottleneck of sustainable development of the city, it will bring a new look of the city in the future and promote the harmonious development of the city. In the next step, the Internet of Things application middleware needs to improve the design of loose integration with the upper application, adapt to sensor equipment more widely, and unify data

standards; the improved data compression algorithm needs to be further improved in the actual environment and system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] C. Lim, G. H. Cho, and J. Kim, "Understanding the linkages of smart-city technologies and applications: key lessons from a text mining approach and a call for future research," *Technological Forecasting and Social Change*, vol. 12, no. 3, pp. 170–192, 2021.

[2] S. Wei, Y. Ma, R. Li, and L. Hu, "Toward smart manufacturing: key technologies and trends driving standardization," *Computer*, vol. 53, no. 4, pp. 46–50, 2020.

[3] Z. Xu, N. Yen, and V. Sugumaran, "Special issue on selected technologies and applications in smart city computing," *Concurrency: Practice and Experience*, vol. 29, no. 24, pp. e4363.1–e4363.2, 2017.

[4] X. Peng, D. Deng, and S. Cheng, "Key technologies of electric power big data and its application prospects in smart grid," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 35, no. 3, pp. 503–511, 2015.

[5] B. Farahani, F. Firouzi, and M. Luecking, "The convergence of IoT and distributed ledger technologies (DLT): opportunities, challenges, and solutions," *Journal of Network and Computer Applications*, vol. 177, no. 2, pp. 102936–102945, 2020.

[6] S. Luo and L. K. Zhang, "The development of International Telecommunication Union (ITU-T) IoT and smart city standardization activities and China's propulsion strategy," *Telecommunications Network Technology*, vol. 4, no. 1, pp. 92–103, 2017.

[7] T. Inui, M. Kohana, S. Okamoto, and M. Kamada, "IoT technologies: state of the art and a software development framework," *Smart Sensors Networks*, vol. 2, pp. 3–18, 2017.

[8] H. T. Cao, "Internet of things and the next generation of smart city," *Mobile Communications*, vol. 9, no. 12, pp. 231–146, 2015.

[9] G. Chen, "Research on key technologies of spatio-temporal information multi-dimensional visualization of smart city," *Mine Surveying*, vol. 12, no. 3, pp. 121–134, 2017.

[10] J. Xie, "Innovation and practice of key technologies for the efficient development of the supergiant Anyue Gas Field," *Natural Gas Industry B*, vol. 7, no. 4, pp. 337–347, 2020.

[11] H. Yi, "Application of key IOT technologies in warehouse automatic identification and management systems," *Logistics Technology*, vol. 13, no. 2, pp. 1321–1334, 2015.

[12] D. Gong and Y. Shi, "Research of key technologies and implementation of campus integrated information service platform," *Intelligent Computer and Applications*, vol. 11, no. 7, pp. 331–345, 2019.

[13] W. Ziguang and W. Ziming, "Research and implementation of some key technologies integrated in WEB information system," *Journal of Henan Science and Technology*, vol. 2, no. 7, pp. 981–992, 2015.

[14] Z. Liu, Y. Liu, B.-J. He, W. Xu, G. Jin, and X. Zhang, "Application and suitability analysis of the key technologies in nearly zero energy buildings in China," *Renewable and Sustainable Energy Reviews*, vol. 101, no. 3, pp. 329–345, 2019.

[15] C. Liu, "Research and application of key technologies of intelligent EMU trains of beijing-zhangjiakou HSR," *China Railway*, vol. 2, no. 7, pp. 245–261, 2019.

[16] X. Gang, X. Dong, and F. Zhu, "Big data technologies and intelligent application system for urban transportation," *Big Data Research*, vol. 3, no. 2, pp. 45–61, 2015.

[17] Q. Zhong, S. U. Zhenmin, and X. Wang, "Key technology applied to construction site of safety monitoring based on integrated application of RFID and BIM," *Building Science*, vol. 3, no. 4, pp. 87–91, 2015.

[18] I. Yutaka, "CEMS and BEMS technologies toward the smart society realization," *Journal of the Society of Instrument and Control Engineers*, vol. 55, no. 7, pp. 604–608, 2016.

[19] P. Huang and E. Coltdwuhan, "The connotation and 10 key technologies of smart manufacturing," *ZTE Technology Journal*, vol. 32, no. 3, pp. 501–513, 2016.

[20] Z.-m. Tong, J.-z. Miao, Y.-s. Li, S.-g. Tong, Q. Zhang, and G.-r. Tan, "Development of electric construction machinery in China: a review of key technologies and future directions," *Journal of Zhejiang University - Science*, vol. 22, no. 4, pp. 245–264, 2021.

[21] S. Hu, S. Tao, and W. Yan, "Advances on enrichment law and key technologies of exploration and development of continental tight oil in China (2016–2018)," *Journal of Natural Gas Geoscience*, vol. 4, no. 6, pp. 87–98, 2019.

[22] Y. Y. Zhang, L. Zhang, and X. U. Qiang, "A research on key technologies of operation and maintenance for smart substation merging unit," *Journal of Jiangxi Vocational and Technical College of Electricity*, vol. 21, no. 3, pp. 43–56, 2017.

[23] S. N. S. Kumar, B. Omar, L. H. Joseph, and N. S. Sivaramakrishnan, "Key measurement principles to strengthen the reliability of loading device technologies: implications to health care practice," " *Polish Annals of Medicine*, vol. 23, no. 2, pp. 123–128, 2016.

[24] S. K. Wang, Y. C. Wang, and Y. Guo, "Research and application of key technologies for fault detection on SLL," *Instrumentation Technology*, vol. 23, pp. 541–553, 2017.

Retraction

Retracted: Design of a Regional Economic Forecasting Model Using Optimal Nonlinear Support Vector Machines

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] T. Zhang, "Design of a Regional Economic Forecasting Model Using Optimal Nonlinear Support Vector Machines," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2900434, 10 pages, 2022.

Hindawi

Research Article

Design of a Regional Economic Forecasting Model Using Optimal Nonlinear Support Vector Machines

Tong Zhang [ID]

Department of Artificial Intelligence, Chongqing College of Finance and Economics, Yongchuan 402160, Chongqing, China

Correspondence should be addressed to Tong Zhang; zhangt766@gmail.com

Forecasting regional economic activity is a progressively significant element of regional economic research. Regional economic prediction can directly assist local, national, and subnational policymakers. Regional economic activity forecast can be employed for defining macroeconomic forces, such as prediction of stock market and cyclicality of national labor market movement. The recent advances of machine learning (ML) models can be employed to solve the time series prediction problem. Since the parameters involved in the ML model considerably influence the performance, the parameter tuning process also becomes essential. With this motivation, this study develops a quasioppositional cuckoo search algorithm (QOCSA) with a nonlinear support vector machine (SVM)-based prediction model, called QOCSO-NLSVM for regional economic prediction. The goal of the QOCSO-NLSVM technique is to identify the present regional economic status. The QOCSO-NLSVM technique has different stages such as clustering, preprocessing, prediction, and optimization. Besides, the QOCSO-NLSVM technique employs the density-based clustering algorithm (DBSCAN) to determine identical states depending upon the per capita NSDP growth trends and socio-economic-demographic features in a state. Moreover, the NLSVM model is employed for the time series prediction process and the parameters involved in it are optimally tuned by the use of the QOCSO algorithm. To showcase the effective performance of the QOCSO-NLSVM technique, a wide range of simulations take place using regional economic data. To determine the current economic situation in a region, the QOCSO-NLSVM technique is used. The simulation results reported the better performance of the QOCSO-NLSVM technique over recent approaches. The QOCSO-NLSVM technique generated effective results with a minimal mean square error of 70.548 or greater. Astonishingly good results were obtained using the QOCSO-NLSVM approach, which had the lowest root mean square error (RMSE) of 8.399.

1. Introduction

The forecasting method predicts future value based on a provided time series data set by making assumptions on future trends and estimating historical data. This is employed for several regions of the decision-making process, like industrial process control, risk management, operations management, demography, and economics [1]. Forecasting is an important problem spanning several domains, involving finance, social science, government, economics, environmental science, politics, medicine, business, and industry. The forecasting problem is categorized as long-term, short-term, and medium-term [2, 3].

Forecasting regional economic activity is an essential component of regional economic study. The regional

economic prediction could directly assist business executives, local, subnational, and national policymakers. These two business executives and policymakers require precise prediction of key economic aggregates, namely, employment, output, and income for medium-long term planning purposes [4]. Regional economic activity forecasts have been employed for explaining macroeconomic forces, involving the cyclicality of national labour market movements and predicting the stock market. Further, multinational agencies and international investors engaged in megaprojects at a regional level also require precise predictions for investment planning reasons [5]. When there is no paucity of research on predicting national economic indicators, the research on regional economic prediction is limited for innovative economies, and in the case of developing nations, zilch [6].

Problems with short-term forecasting are those that deal with predicting events in a shorter period of time (months, days, and weeks). Forecasting concerns could go much beyond 1-2 years into the future, with medium-term forecasts extending into the future as well.

The forecasting method connected to economic problems is utilized for predicting economic variables in several countries. The industry volatility prediction, critical to several important problems in business [7], and the prediction of the unemployment rates that define the country's economic and social development [8, 9]. Radial basis function networks (RBF) and backpropagation are the ANN architectures that are used in economic fields. The artificial neural networks (ANN) technique was broadly examined in economic analysis. The ANN is a computation system that is performed in hardware or software under the effect of biological studies about the human brain. Several authors admit that the ANN method is the better performing nonlinear analysis technique as well as one of the best predictors [10]. The ANN architecture employed in economic fields is radial basis function networks (RBF) and backpropagation.

This study designs a quasioppositional cuckoo search algorithm (QOCSA) with a nonlinear support vector machine (SVM)-based prediction model, called QOCSO-NLSVM for regional economic prediction. The QOCSO-NLSVM technique involves the design of the density-based clustering algorithm (DBSCAN) to determine the identical states depending upon the per capita NSDP growth trends and socioeconomic-demographic features in a state. Besides, the NLSVM model is elected for the time series prediction process and the parameters involved in it are optimally tuned by the use of the QOCSO algorithm. The experimental validation of the QOCSO-NLSVM technique and the results are examined in various aspects.

The rest of the research work is organized as follows. Section 2 provides the recently developed techniques, Section 3 elaborates the QOCSO-NLSVM technique. Then, Section 4 provides the performance validation, and Section 5 concludes the outcomes of the research.

2. Literature Review

Mishra and Ayyub [11] introduced a DL architecture in which the hierarchical clustering analysis (HCA) is utilized for predicting growth. The presented method comprises HCA and DTW techniques that are initially applied for identifying similar socio-economic-demographic features within a provided state and similar states according to per capita NSDP growth trends, to create a fine-tuned training dataset for predicting all the states' NSDP per capita growth. Lv et al. [12] developed a LightGBM-enhanced LSTM for realizing stock price prediction, and LSTM is utilized for predicting the Shenzhen and Shanghai 300 indexes, respectively. The simulation result shows that the LightGBM-LSTM has a better capacity for tracking stock index price trends and the maximum prediction performance, and its effects are superior to the RNN and GRU methods. LightGBM-optimized LSTM for short-term stock price

forecasting. To compare its performance with other deep network models such as RNN (recurrent neural network) and GRU (gated recurrent unit), the LightGBM-LSTM, RNN, and GRU are used to predict the Shanghai and Shenzhen 300 indexes, respectively. Experiment results demonstrate that the LightGBM-LSTM has the highest prediction accuracy and the best ability to track stock index price trends.

Zhu et al. [13] designed an experiment whose samples originated from information on 7 quoted core enterprises (CEs) and 46 quoted SMEs in the Chinese security markets. Matta et al. [14] introduced a relative assessment of various prediction techniques using the Gaussian process regression and ANN methods (MLP and RBFNN). Two real-time datasets were utilized for evaluating the prediction method presented in the study. These datasets were normalized to values amongst one and zero. Next, the data training was implemented and, when it was constructed, a system was utilized for generating the predictions. Therefore, observations were made to validate how precisely the fitted method predicts the values.

Chatzis et al. [15] integrated distinct ML methods that were proposed with daily currency, stock, and bond data from thirty-nine countries that cover a larger spectrum of economies. It especially leverages the advantages of a sequence of techniques that includes Classifier Trees, SVM, NN, RF, XGBoost, and DNN. Sun et al. [16] verified the cointegration relationships and Granger causality between tourist arrivals in Beijing and the internet search index. This experiment result suggests that compared to standard methods, the presented KELM model that incorporates tourist volume series with Google and Baidu Index could significantly enhance the prediction performances in terms of robustness analysis and forecasting accuracy.

3. The Proposed Model

In this study, an effective QOCSO-NLSVM technique has been developed for regional economic prediction. The QOCSO-NLSVM technique encompasses several subprocesses, namely, DTW-based preprocessing, DBSCAN-based clustering, NLSVM-based prediction, and QOCSO-based parameter optimization. Figure 1 illustrates the overall working process of the QOCSO-NLSVM technique.

3.1. Data Preprocessing. One of the primary methods used to capture similarities among two regions, or among pairs of factors within a provided region according to time-series data is named dynamic time warping (DTW). DTW is an effective method utilized for learning similarity based on distance between two sequences that might differ in speed and quantifying time-based similarities among any two pairs. Generally, DTW is an ML method which estimates an optimum match between two provided sequences with some restrictions. The sequence is "warped" nonlinearly in the time dimension to define measures of their similarity, independent of nonlinear variation in the time dimension. The Euclidean distance uses the distance among every pair of the
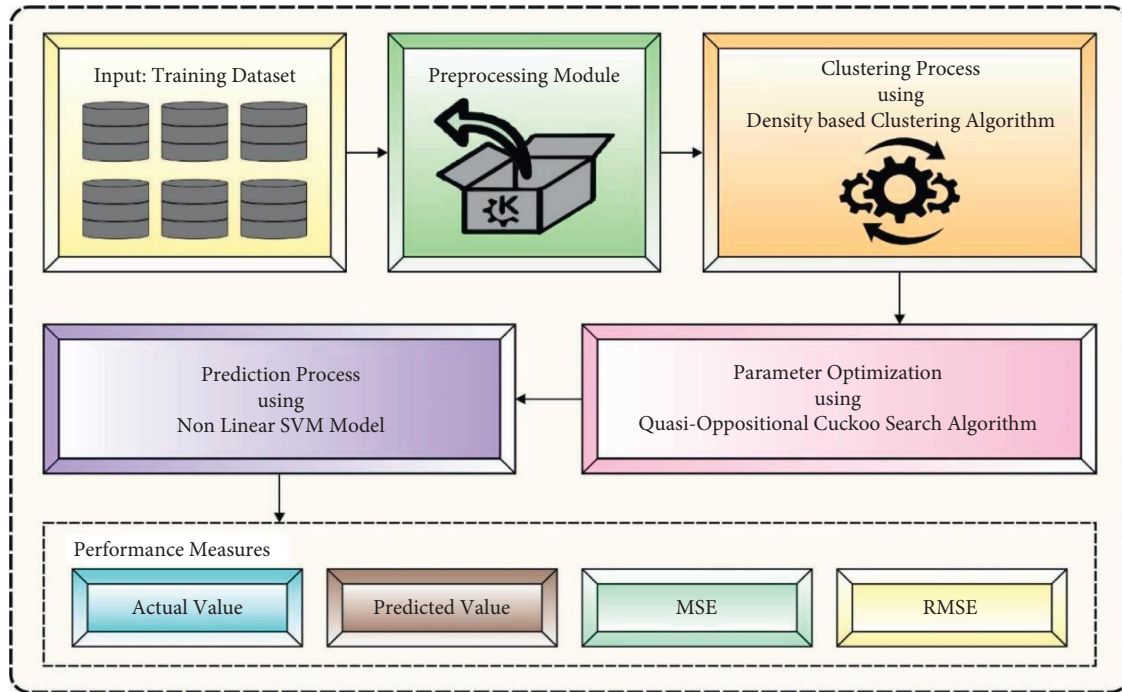
FIGURE 1: System architecture of the QOCSO-NLSVM method.

time series and compares it with the Euclidean distance. Simultaneously, the DTW searches for optimal alignments among the two-time series. Furthermore, all the points are utilized for comparing the points to make the best possible alignments among the two-time series according to their distance matrix.

3.2. Process Involved in the DBSCAN Technique.

DBSCAN might find distinct clusters based on the assessed density distribution. It could recognise structured groupings without knowing their numbers. The following illustrates DBSCAN's basic premise: DBSCAN finds each point in the neighbourhood of a random unvisited point p, where it denotes the neighbourhood's maximum radius from p. To construct a dense zone, MinPts is the minimum number of points required. When MinPts is in the distance, p denotes a core point. When p is a core point, all points in its vicinity are grouped together. DBSCAN detects each density-reachable point in the cluster and adds it to a comparable cluster. When a point q is densely approachable from other core points but its neighbourhood is less than MinPts, it is a border point. An outlier or noisy point is one that is not accessible from other locations. DBSCAN achieves clustering by extracting clusters consecutively. Rep until no more density-reachable points are identified, and the final cluster is reached. DBSCAN divides a set of points into low-noise border points and high-density. The purpose of DBSCAN is to identify identical states based on a state's per capita NSDP growth trends and socioeconomic-demographic characteristics. DBSCAN was capable of detecting a variety of clusters based on the density distribution that was assessed. The DBSCAN methodology permits the calculation of identical states based on per capita income.

Assume two points x and y, $d(x, y)$ represent the similarities among them, $\Gamma\varepsilon(x)$ denotes the ε-neighbourhood of x, in which $\Gamma\varepsilon(x) = \{y \in V | d(x, y) \leq \varepsilon\} \cdot \rho(x) = |\Gamma\varepsilon(x)|$ indicates the density value of x:

$$S(x) = \begin{cases} 1, & \text{core point with } \rho(x) \geq \text{MinPts}, \\ 0, & \text{border point with } 1 < \rho(x) < \text{MinPts}, \\ -1, & \text{noise with } \rho(x) = 1. \end{cases} \quad (1)$$

3.3. Structure of the NLSVM Model.

During the prediction process, the NLSVM model receives the clustered data as input to predict the output. Assume a trained set $\{x_k, y_k\}_{k=1}^{N}$ with input data $x_k \in \mathbb{R}^n$ and respective binary class label $y_k \in \{-1, +1\}$, the SVM classification initiates from the subsequent assumption:

$$\begin{cases} w^T\varphi(x_k) + b \geq +1, & \text{if } y_k = +1, \\ w^T\varphi(x_k) + b \leq -1, & \text{if } y_k = -1. \end{cases} \quad (2)$$

That is equal to

$$y_k[w^T\varphi(x_k) + b] \geq 1, \quad k = 1, \ldots, N. \quad (3)$$

Now, the nonlinear function $\varphi(\cdot): \mathbb{R}^n \longrightarrow \mathbb{R}^{n_h}$ maps the input space to a high-dimensional feature space. It is noteworthy that the n_h dimension of this space is determined in an implicit manner (it is an infinite dimension). The b represent a bias as follows:

$$y(x) = \text{sign}[w^T\varphi(x) + b]. \quad (4)$$

But, at the same time, it is never evaluated in this form. One determines the optimization issue:

$$\min \mathscr{J}(w, \xi) = \frac{1}{2} w^T w + c \sum_{k=1}^{N} \xi_k, \qquad (5)$$

subjected to

$$\begin{cases} y_k \left[w^T \varphi(x_k) + b \right] \geq 1 - \xi_k, & k = 1, \ldots, N, \\ \xi_k \geq 0, & k = 1, \ldots, N. \end{cases} \qquad (6)$$

To permit misclassification in the subset of inequalities (because of overlapping distribution), the minimalization of $\|w\|_2$ corresponds to a maximalization of the margin among the two classes. c indicates a positive real constant and must be taken into account as a tuning parameter. The Lagrangian can be expressed as follows [17]:

$$L(w, b, \xi; \alpha, \nu) = \mathscr{J}(w, \xi) - \sum_{k=1}^{N} \alpha_k$$

$$\left\{ y_k \left[w^T \varphi(x_k) + b \right] - 1 + \xi_k \right\} - \sum_{k=1}^{N} \nu_k \xi_k. \qquad (7)$$

The Lagrange multiplier is $\alpha_k \geq 0, \nu_k \geq 0, (k = 1, N)$. Figure 2 depicts the SVM hyperplane. It is familiar from the optimization concept that the solutions are considered by the saddle points of the Lagrangian:

$$\max_{\alpha, \nu} \min_{w, b, \xi} \mathscr{L}(w, b, \xi; \alpha, \nu). \qquad (8)$$

One attains

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \longrightarrow w = \displaystyle\sum_{k=1}^{N} \alpha_k y_k \varphi(x_k), \\ \\ \dfrac{\partial L}{\partial b} = 0 \longrightarrow \displaystyle\sum_{k=1}^{N} \alpha_k y_k = 0, \\ \\ \dfrac{\partial L}{\partial \xi_k} = 0 \longrightarrow 0 \leq \alpha_k \leq c, \quad k = 1, \ldots, N. \end{cases} \qquad (9)$$

By substituting w in the Lagrangian, one attains the subsequent binary problems (in the Lagrange multiplier α), i.e., the quadratic programming problems:

$$\max_{\alpha} \mathcal{Q}(\alpha) = -\frac{1}{2} \sum_{k,l=1}^{N} y_k y_l K(x_k, x_l) \alpha_k \alpha_l + \sum_{k=1}^{N} \alpha_k. \qquad (10)$$

Thus,

$$\begin{cases} \displaystyle\sum_{k=1}^{N} \alpha_k y_k = 0, \\ 0 \leq \alpha_k \leq c, \quad k = 1, \ldots, N. \end{cases} \qquad (11)$$

Now w and $\varphi(x_k)$ are not estimated. According to the Mercer condition, one takes a kernel as

$$K(x_k, x_l) = \varphi(x_k)^T \varphi(x_l). \qquad (12)$$
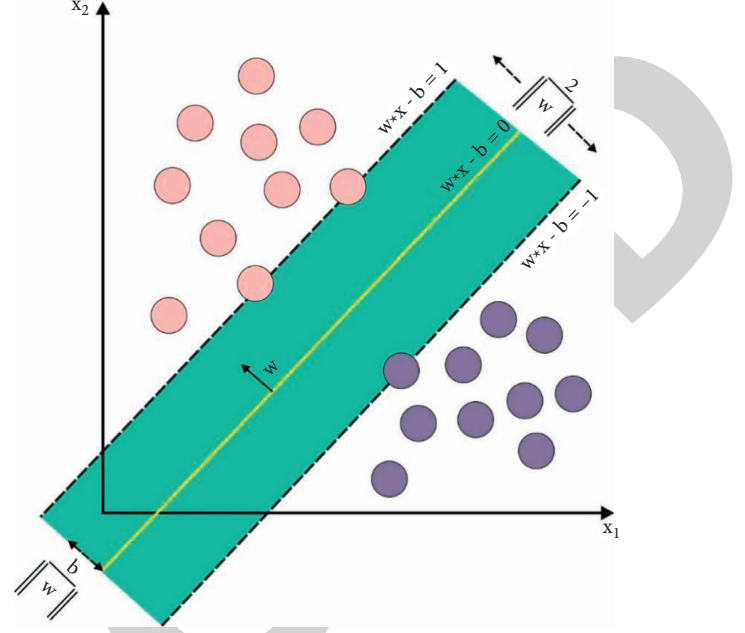
Figure 2: SVM hyperplane.

Lastly, in binary space, the nonlinear SVM classifiers become

$$y(x) = \text{sign} \left[\sum_{k=1}^{N} \alpha_k y_k K(x, x_k) + b \right]. \qquad (13)$$

α_k is a positive real constant, and b is a real constant. The nonzero Lagrange multiplier α_k is known as support value. The respective data point is known as a support vector and is placed near the decision boundary. This is the data point that contributes to the classification method. The bias b follows from the KKT condition that isn't considered further.

Various selections for the kernel $K(\cdot, \cdot)$ are feasible.

(i) $K(x, x_k) = x_k^T x$ (linear SVM)
(ii) $K(x, x_k) = (x_k^T x + 1)^d$ (polynomial SVM of degree d)
(iii) $K(x, x_k) = \exp\left\{ -\|x - x_k\|_2^2 / \sigma^2 \right\}$ (RBF kemel)
(iv) $K(x, x_k) = \tanh(\kappa x_k^T x + \theta)$ (MLP SVM)

The Mercer conditions hold for each σ value in the RBF case, but not for each feasible selection of κ, θ in the MLP case. In the case of an MLP or RBF kernel, the amount of hidden units corresponds to the number of support vectors.

3.4. Design of the QOCSO Algorithm for Parameter Tuning.

For optimally tuning the weight values of the NLSVM model, the QOCSO algorithm is utilized. The CSO algorithm is assumed as a metaheuristic technique that was primarily established by Yang and Deb [18]. Actually, this CSO method simulates the breeding performance of cuckoo birds that are supposed to be a type of parasitism. The cuckoo birds place their eggs from other nests and play to host the egg. The cuckoo birds attempt for raising the hatch possibility of their individual eggs by generating them the same as

```
Begin.
    Objective functions of f(x), x = (x1, x2, ..., xd)T
    Populations initialization of n host nests xi,
    While (t < Maximum_iteration) or (termination condition)
        Get a cuckoo arbitrarily via Lévy flight
        Determine the qualities/fitness as Fi
        Select nest amongst (n, j) arbitrarily
    If (Fi ≥ Fj),
        Substitute j in newly attained solution;
    End
        An fraction (pa), poor nests are discarded and new one is derived;
        Retain optimal solution, (with quality solution);
        Sort the solutions and determine current best
    End
    Postprocess and visualize results
End
```

ALGORITHM 1: Pseudocode of the CSO algorithm.

TABLE 1: Actual and predicted analysis of the QOCSO-NLSVM technique with varying years.

Years	Actuals	Run: 1	Run: 2	Run: 3	Run: 4	Run: 5
2012	16855.712	16854.72	16784.71	16716.72	16799.72	16851.73
2013	17037.957	17134.99	17056.94	17098.97	17102.96	17028.95
2014	17675.816	17767.80	17563.80	17820.81	17781.81	17734.82
2015	18176.991	18110.02	18125.98	18285.98	18211.01	18294.03
2016	18860.411	18924.41	18782.41	18806.42	18750.44	18970.43
2017	19999.445	19966.48	20106.45	19960.45	19859.44	19874.43
2018	21138.478	21028.50	21282.49	21097.47	21184.48	21098.48
2019	21730.776	21822.77	21680.80	21651.76	21680.76	21627.80
2020	22231.951	22310.96	22177.95	22379.96	22330.96	22374.97

the host egg with respect to size, shape, and colour, or by throwing other native eggs (Algorithm 1).

In the CSO technique, cuckoo eggs from distinct nests signify the generation of candidate solutions to optimize problems. Actually, the search starts with particular nests with a solution per nest. This solution was progressed dependent upon the model of cuckoo's recognition (p) which was inspired by eliminating the solution of exchanging novel ones.

In the CSO method, a random walk was utilized dependent upon the Lévy flight distribution for producing novel candidate solutions (cuckoos) in the present one as follows:

$$\text{cuckoo}_i^{(t+1)} = \text{cuckoo}_i^{(t)} + a \oplus \text{Levy}(\lambda), \quad (14)$$

where $\text{cuckoo}_i^{(t+1)}$ refers the ith cuckoo value t. An a and λ stand for step sizes (generally fixed to one) and coefficients ($1 < \lambda < 3$) correspondingly. A number of novel solutions were created in the optimum present ones by Lévy walks for performing a local search with self-improvement [19]. Besides, a few novel solutions were created away from the optimum present ones. This reduces the chance of getting stuck from the local minimal and ensures the searching ability. The CS execution also makes sure elitism as the optimal nest is retained under the iteration.

The OBL method was proposed with the aim of decreasing the computation time and improving the ability of various EAs [20]. Therefore, the comparisons among an arbitrary CSO algorithm and its opposite might result in the global optimal with fast convergence rates. Further, the quasiopposite number and showed that it is nearer to the optimum solution when compared to the opposite number. Therefore, the population initialization of this method is created according to the QOBL concept. For arbitrary number $\chi \in [a, b]$, its opposite number χ_0 is represented as follows:

$$x_0 = a + b - x. \quad (15)$$

However, the opposite point for multidimensional searching space (dimension) is determined by the following equation:

$$x_0^i = a^j + b^i - x^i, \quad i = 1, 2, \ldots, d. \quad (16)$$

The quasiopposite no. x_{qo} of arbitrary no. $\chi \in [a, b]$ is represented as follows [21]:

$$x_{qo} = \text{rand}\left(\frac{a+b}{2}, x_0\right). \quad (17)$$

Likewise, the quasiopposite point for multidimensional searching space (d dimension) is determined by the following equation:
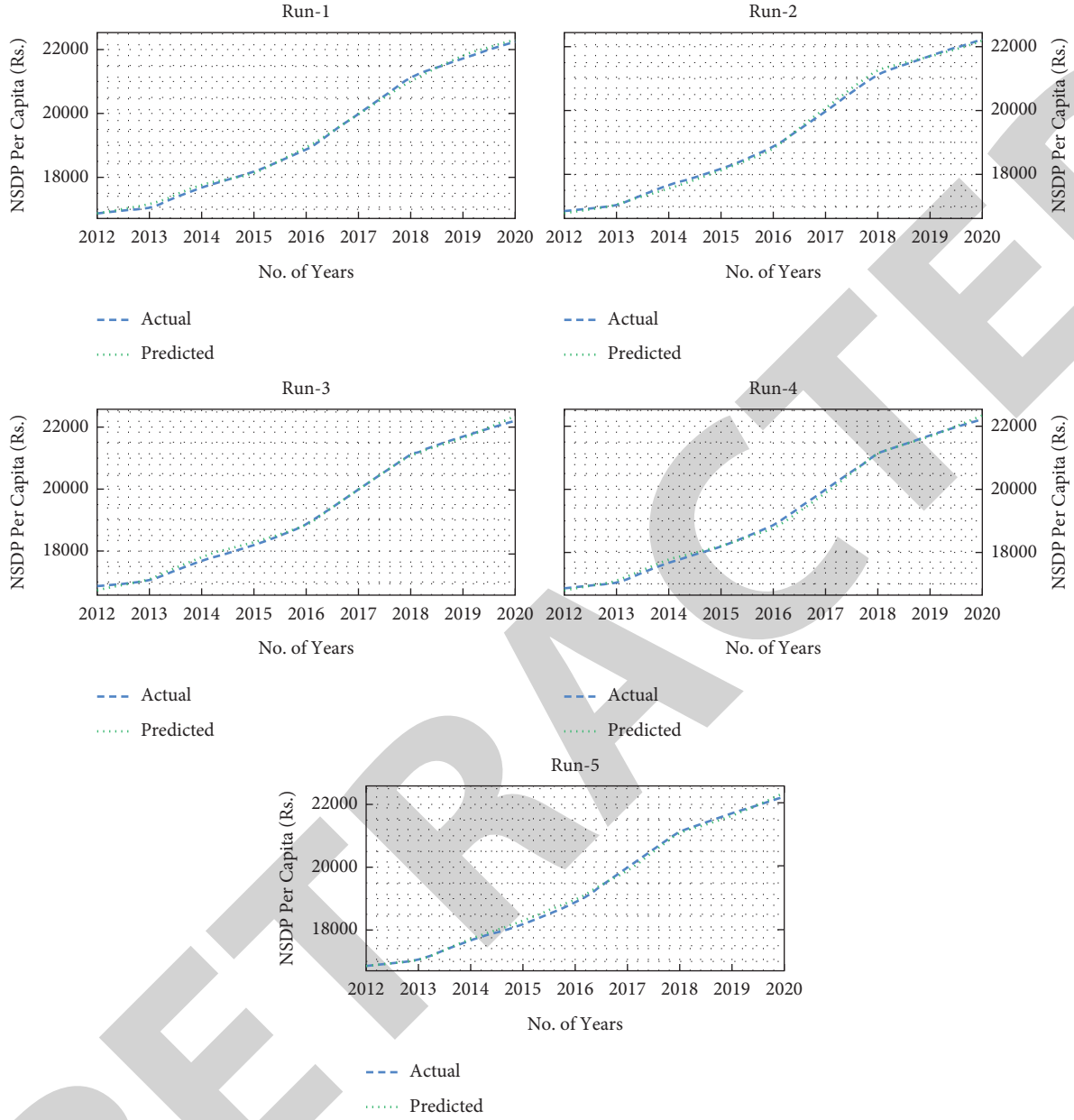
Figure 3: Actual and prediction analysis of the QOCSO-NLSVM technique.

$$x_{qo}^i = \text{rand}\left(\frac{a^i + b^i}{2}, x_0^i\right). \qquad (18)$$

For obtaining an objective function that could generalize the SVM outcome with no utilization of testing data, the cross validation approach is utilized. The cross validation process partitions the training dataset D randomly into S different parts $\{G_s, \quad s = 1, \ldots, S\}$, and utilizes $(S-1)$ parts to train the model and to test the model. This process gets iterated for S times by varying the lasting parts, and the generalization efficiency can be determined by the use of MSE (mean squared error) over every test result.

$$MSE_{CV} = \frac{1}{N} \sum_{s=1}^{S} \sum_{i \in G_s} (y_i - f(x_i|\theta_s))^2, \qquad (19)$$

where G_s indicates the s-th part for the testing process and θ_s signifies the solution vector attained at the time of training process.

4. Performance Evaluation and Discussion

The performance validation of the QOCSO-NLSVM technique using the economic data from the Niti Aayog website and the Reserve Bank of India were inspected. The data
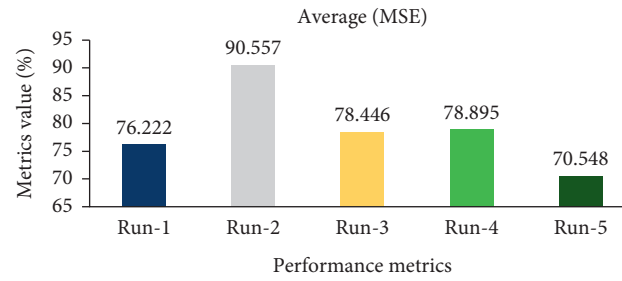
Figure 4: MSE analysis of the QOCSO-NLSVM technique with distinct runs.

includes several features such as fiscal deficits, revenue deficits, interest payments, capital expenditure, nominal NSDP series, social sector expenditure, electricity generation, infrastructure projects, per capita NSDP at factor cost (at constant prices), per capita NSDP, number of factories, state-wise fixed capital, sectoral growth rate, and pattern of land use. Table 1 and Figure 3 investigate the actual and predicted result analysis of the QOCSO-NLSVM technique over distinct years. The results portrayed that the QOCSO-NLSVM technique predicted the economic status much closer to the actual value under all runs.

For instance, with an actual value of 16855.712, the QOCSO-NLSVM technique has attained predicted values of 16854.72, 16784.71, 16716.72, 16799.72, and 16851.73 under runs 1–5, respectively. At the same time, with the actual values of 18176.991, the QOCSO-NLSVM system has accomplished forecasted values of 18110.02, 18125.98, 18285.98, 18211.01, and 18294.03 under runs 1–5 correspondingly. Furthermore, with the actual values of 21138.478, the QOCSO-NLSVM method has achieved forecasted values of 21028.50, 21282.49, 21097.47, 21184.48, and 21098.48 under runs 1–5 correspondingly. Moreover, with the actual values of 22231.951, the QOCSO-NLSVM algorithm has reached predicted values of 22310.96, 22177.95, 22379.96, 22330.96, and 22374.97 under runs 1–5 correspondingly.

A brief MSE analysis of the QOCSO-NLSVM technique under various runs and years is provided in Figure 4 and Table 2. The experimental values are denoted by the QOCSO-NLSVM technique, which has resulted in an effective outcome with minimal MSE values. For instance, in the year 2012, the QOCSO-NLSVM technique resulted in at least MSE of 70.997, 138.996, 55.994, 3.979, and 0.995, respectively. Simultaneously, in the year 2015, the QOCSO-NLSVM system has resulted in a minimum MSE of 51.009, 108.987, 34.014, 117.037, and 66.972 correspondingly. Simultaneously, in the year 2018, the QOCSO-NLSVM model has resulted in a minimum MSE of 144.010, 41.006, 46.005, 39.996, and 109.978 correspondingly. Likewise, in the year 2020, the QOCSO-NLSVM method has resulted in a minimum MSE of 54.003, 148.011, 99.014, 143.021, and 79.009 correspondingly.

A brief RMSE analysis of the QOCSO-NLSVM method over many years and runs has been demonstrated in Table 3 and Figure 5. The experiment values showed that the QOCSO-NLSVM method has resulted in outstanding results with the smallest RMSE value. For example, in the year

Table 2: MSE analysis of the QOCSO-NLSVM technique with distinct runs.

Years	Run: 1	Run: 2	Run: 3	Run: 4	Run: 5
2012	70.997	138.996	55.994	3.979	0.995
2013	18.987	61.017	65.000	9.006	97.028
2014	112.011	144.993	105.997	59.004	91.982
2015	51.009	108.987	34.014	117.037	66.972
2016	77.999	53.994	109.975	110.019	64.000
2017	107.001	38.997	140.001	125.018	32.967
2018	144.010	41.006	46.005	39.996	109.978
2019	49.977	79.014	50.013	102.978	91.997
2020	54.003	148.011	99.014	143.021	79.009
Average	76.222	90.557	78.446	78.895	70.548

2012, the QOCSO-NLSVM system resulted in a minimal RMSE of 8.426, 22.790, 7.483, 1.995, and 0.997 correspondingly. Concurrently, in the year 2015, the QOCSO-NLSVM approach resulted in a minimum RMSE of 7.142, 10.440, 5.832, 10.818, and 8.184 correspondingly. Simultaneously, in the year 2018, the QOCSO-NLSVM process has resulted in the smallest RMSE of 12.000, 6.404, 6.783, 6.324, and 10.487 correspondingly. Likewise, in the year 2020, the QOCSO-NLSVM method has resulted in a minimal RMSE of 7.349, 12.166, 9.951, 11.959, and 8.889 correspondingly.

Table 4 presents a full comparison study of the QOCSO-NLSVM approach.

Figure 6 offers the MSE analysis of the QOCSO-NLSVM technique with recent methods. The figure shows that the LSTM and ARIMA models have obtained poor performance with a higher MSE of 149.997 and 142.235, respectively. Similarly, the GRU and multivariate LSTM models reached a moderate MSE of 128.357 and 95.184, respectively. However, the QOCSO-NLSVM technique has accomplished effective outcomes with a minimal MSE of 70.548.

Figure 7 provides the RMSE of the QOCSO-NLSVM model with current methodologies. The abovementioned figure exhibits that the ARIMA and LSTM systems have gained poor performance with a high RMSE of 11.926 and 12.247 correspondingly. Simultaneously, the multivariate LSTM and GRU methods have attained reasonable RMSE of 9.756 and 11.329, respectively. But, the QOCSO-NLSVM process has gained remarkable results with the smallest RMSE of 8.399.

TABLE 3: RMSE analysis of the QOCSO-NLSVM technique with distinct runs.

Years	Run: 1	Run: 2	Run: 3	Run: 4	Run: 5
2012	8.426	11.790	7.483	1.995	0.997
2013	4.357	7.811	8.062	3.001	9.850
2014	10.584	12.041	10.296	7.681	9.591
2015	7.142	10.440	5.832	10.818	8.184
2016	8.832	7.348	10.487	10.489	8.000
2017	10.344	6.245	11.832	11.181	5.742
2018	12.000	6.404	6.783	6.324	10.487
2019	7.069	8.889	7.072	10.148	9.592
2020	7.349	12.166	9.951	11.959	8.889
Average	8.456	9.237	8.644	8.177	7.926
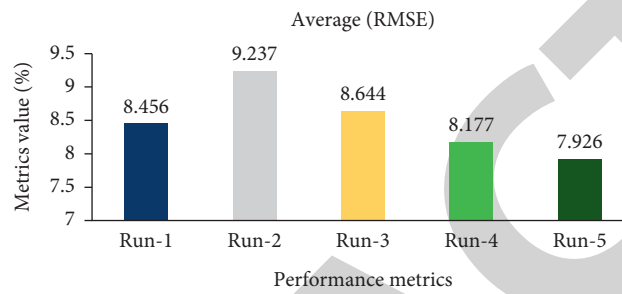
FIGURE 5: RMSE analysis of the QOCSO-NLSVM technique with distinct runs.

TABLE 4: Comparative analysis of the QOCSO-NLSVM technique with existing approaches.

	MSE	RMSE
LSTM	149.997	12.247
ARIMA	142.235	11.926
GRU	128.357	11.329
Multivariate LSTM	095.184	09.756
QOCSO-NLSVM	070.548	08.399
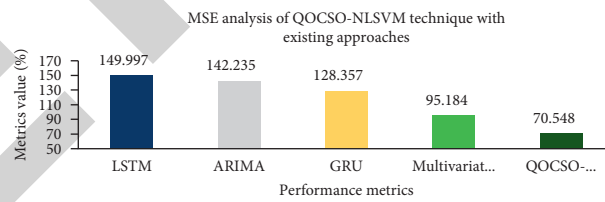
FIGURE 6: MSE analysis of the QOCSO-NLSVM technique with existing approaches.
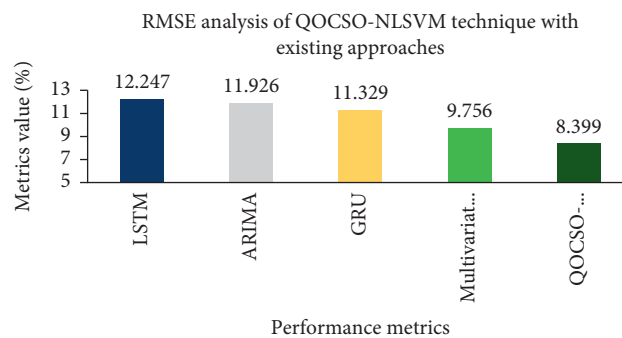
FIGURE 7: RMSE analysis of the QOCSO-NLSVM technique with existing approaches.

From the abovementioned figures, it is ensured that the QOCSO-NLSVM model is an effective regional economic prediction method over the other existing techniques.

5. Conclusion

In this research, a proposed QOCSO-NLSVM technique has been developed for regional economic prediction. The QOCSO-NLSVM technique encompasses several subprocesses, namely, DTW based preprocessing, DBSCAN-based clustering, NLSVM-based prediction, and QOCSO-based parameter optimization. The use of the DBSCAN model enables the computation of identical states depending upon the per capita NSDP growth trends and socioeconomic-demographic features in a state. In addition, the application of the QOCSO algorithm helps to properly select the parameter values and thereby reaches the maximum predictive outcomes. The QOCSO-NLSVM technique is used to discover identical states based on per capita NSDP growth trends and socioeconomic-demographic characteristics in a state. QOCSO-NLSVM is used to run a variety of simulations on regional economic data and is also used to assess a region's present economic position. The experimental validation of the QOCSO-NLSVM technique and the results are examined in various aspects. The comparative analysis revealed the enhanced outcomes of the QOCSO-NLSVM technique over the recent approaches. With a minimum MSE of 70.548, the QOCSO-NLSVM approach produced effective results. The QOCSO-NLSVM technique had remarkable results, achieving the lowest root mean square error (RMSE) of 8.399. In the future, advanced DL models can be used to improve the overall prediction outcomes.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest with any financial organizations regarding the material reported in this manuscript.

References

[1] C. H. dos Santos, R. D. C. Lima, F. Leal, J. A. de Queiroz, P. P. Balestrassi, and J. A. B. Montevechi, "A decision support tool for operational planning: a Digital Twin using simulation and forecasting methods," *Production*, vol. 30, Article ID e20200018, 2020.

[2] D. C. Montgomery, C. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, Wiley, Hoboken, NJ, USA, 2015.

[3] R. Lehmann and K. Wohlrabe, "Regional Economic Forecasting: State-Of-The-Art Methodology and Future Challenges," *Economics and Business Letters*, vol. 3, no. 4, pp. 218–231, 2014.

[4] N. Gonzalez Prieto, P. Loungani, and S. Mishra, "What lies beneath? A sub-national look at okun's law in the United States," *Open Economies Review*, vol. 29, no. 4, pp. 835–852, 2018.

[5] D. S. Rickman, "Regional science research and the practice of regional economic forecasting: less is not more," in *Regional Research Frontiers*, vol. 1, pp. 135–149, Springer, Cham. New York, USA, 2017.

[6] W. Puchalsky, G. T. Ribeiro, C. P. da Veiga, R. Z. Freire, and L. d. Santos Coelho, "Agribusiness time series forecasting using Wavelet neural networks and metaheuristic optimization: an analysis of the soybean sack price and perishable products demand," *International Journal of Production Economics*, vol. 203, pp. 174–189, 2018.

[7] S. Tuo, T. Chen, H. He et al., "A regional industrial economic forecasting model based on a deep convolutional neural network and big data," *Sustainability*, vol. 13, no. 22, Article ID 12789, 2021.

[8] S. Gu, B. Kelly, and D. Xiu, "Empirical asset pricing via machine learning," *Review of Financial Studies*, vol. 33, no. 5, pp. 2223–2273, 2020.

[9] M. Hadjicharalambous, M. M. Polycarpou, and C. G. Panayiotou, "Neural network-based construction of online prediction intervals," *Neural Computing & Applications*, vol. 32, no. 11, pp. 6715–6733, 2020.

[10] N. Dritsakis and P. Klazoglou, "Forecasting unemployment rates in USA using Box-Jenkins methodology," *International Journal of Economics and Financial Issues*, vol. 8, no. 1, pp. 9–20, 2018.

[11] S. Mishra and B. M. Ayyub, "Predicting regional economic activity using artificial intelligence (AI) methods: case study with Indian states," *IASSI-Quarterly*, vol. 37, no. 3and4, pp. 330–361, 2018.

[12] J. Lv, C. Wang, W. Gao, and Q. Zhao, "An Economic Forecasting Method Based on the LightGBM-Optimized LSTM and Time-Series Model," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8128879, 10 pages, 2021.

[13] Y. Zhu, L. Zhou, C. Xie, G.-J. Wang, and T. V. Nguyen, "Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach," *International Journal of Production Economics*, vol. 211, pp. 22–33, 2019.

[14] C. E. D. Matta, N. M. P. Bianchesi, M. S. D. Oliveira, P. P. Balestrassi, and F. Leal, "A comparative study of forecasting methods using real-life econometric series data," *Production*, vol. 31, 2021.

[15] S. P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, and N. Vlachogiannakis, "Forecasting stock market crisis events using deep and statistical machine learning techniques," *Expert Systems with Applications*, vol. 112, pp. 353–371, 2018.

[16] S. Sun, Y. Wei, K.-L. Tsui, and S. Wang, "Forecasting tourist arrivals with machine learning and internet search index," *Tourism Management*, vol. 70, pp. 1–10, 2019.

[17] J. A. Suykens, "Nonlinear modelling and support vector machines,"vol. 1, pp. 287–294, in *Proceedings of the IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering measurement in the age of informatics (Cat. No. 01CH 37188)*, vol. 1, IEEE, Budapest, Hungary, 2001, May.

[18] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *Proceedings of the 2009 World Congress on Nature & Biologically Inspired Computing*, pp. 210–214, IEEE, Coimbatore, India, December, 2009.

[19] D. Paulraj, "A gradient boosted decision tree-based sentiment classification of twitter data," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 18, no. 4, Article ID 20500277, 2020.

Research Article

Load Balancing Cloud Storage Data Distribution Strategy of Internet of Things Terminal Nodes considering Access Cost

Jiansheng Wu ⓘ,[1] Weimin Xu ⓘ,[1] and Jiarong Xia ⓘ[2]

[1]*China Tobacco Zhejiang Industrial Co., Ltd.,, Ningbo, Zhejiang 315504, China*
[2]*School of Mathmatics, Hangzhou Normal University, Hangzhou 311121, China*

Correspondence should be addressed to Jiarong Xia; jrxia@hznu.edu.cn

With the rapid development of Internet of Things (IoT) technology, IoT terminal nodes are facing many challenges in data storage, distribution, and data management. In particular, in the IoT terminal nodes considering access cost, the corresponding data distribution and storage are professional, complex, and miscellaneous. Based on the abovementioned current situation, this article innovatively proposes a complex sensor data placement algorithm based on the cloud storage distribution of IoT terminal nodes. Under this algorithm, the accurate division of IoT data I/O methods is realized through reasonable configuration. Through the adaptive sensing algorithm, while fully considering the access cost of the algorithm, the performance of the IoT data storage system is further optimized. In the corresponding terminal node load balancing problem, this article innovatively proposes the terminal node data sorting and distribution algorithm through the node data. The sorting and distribution algorithm realizes the precise segmentation of the IoT data to be processed, thereby realizing the improvement of data reading and processing speed. Based on the proposed algorithm, this article designs a load balancing cloud storage data distribution optimization system of IoT terminal nodes considering access cost and carries out experimental verification in a real environment. The experimental results show that the data pattern division accuracy corresponding to the proposed distribution strategy is improved to 97.13% and the corresponding data access efficiency is improved to 98.3%, compared with the traditional distribution strategy. Therefore, the data distribution strategy proposed in this article has obvious performance advantages and further promotion value.

1. Introduction

The main architecture of the Internet of Things (IoT) includes the perception layer, network layer, management layer and application layer. The conventional sensing layer is mainly used to collect relevant physical information, such as temperature, humidity, air composition, and optical signals. The continuous collection and storage of data by the relevant sensor networks of the IoT cause problems of data distribution and data management [1–3]. The relevant data of the IoT are mainly classified into structured data and unstructured data. The corresponding structured data mainly depend on open-source databases and other products, while the storage of the corresponding unstructured data depends on the database software developed by the corresponding unstructured data model. At the data storage level of the IoT, cloud storage technology, as the development direction of

data storage, mainly processes and analyzes a large number of different types of data generated by the IoT through cluster application, network technology, and distributed storage technology and provides external data access and storage [4, 5]. The conventional cloud storage architecture includes the application interface layer, infrastructure layer, and storage layer. The corresponding application interface layer mainly includes network access, user authentication, permission management, public API interface, application software interface, and network service interface. Moreover, the corresponding infrastructure layer includes the cluster system, distributed file system, grid system, content distribution and data deduplication, data compression, and data encryption processing. The corresponding storage layer mainly includes storage virtualization, centralized storage management, status monitoring, maintenance, and upgrading [6]. Data placement and node load balancing are

important technologies of distributed storage in cloud storage technology. Load balancing mainly realizes the dynamic expansion of storage resources to further improve the utilization of resources, minimize the data response time, and further reduce the fault recovery and data switching time to ensure the quality of user data access in the IoT system [7, 8]. At the level of corresponding data placement and distribution, the flat distribution of the data system is of great significance. The directory organization corresponding to the traditional data distribution is unreasonable, and the node data distribution is too scattered or too centralized, which will further affect the response-ability of the data system, further reduce the system performance, and increase the data maintenance cost [9, 10]. Compared with other storage models, cloud storage has the following advantages: (1) easy to expand: the storage space is expanded in time according to the number of users and space of the server, which will also not affect the use of the front-end users; (2) reliable and safe: data synchronization effectively avoids media storage of data. Cloud storage can fully realize load balancing, which can realize the flow distribution control service of distributing access traffic to multiple back-end cloud servers according to the forwarding strategy. More-over, the external service capabilities of the application system are extended through traffic distribution, and the availability of the application system is improved by eliminating single points of failure.

The traditional cloud storage data distribution strategy of load balancing of IoT terminal nodes has some disadvantages in data placement and load balancing technology. At the data placement level, the disadvantages are as follows: differences in data storage performance and availability storage cost between different data storage service providers and differences in different data reading and writing modes in the IoT sensor network. Therefore, the traditional data placement algorithm only meets a single optimization goal and cannot meet the application storage requirements under the huge amount of data in IoT. Based on this, constructing an efficient data placement algorithm to optimize the performance of different data access is of great significance to improve the performance of cloud storage data systems for load balancing of IoT terminals [11, 12]. For the problem of load balancing, after the corresponding IoT data are initialized, as the storage system continues to serve various big data-related applications on the upper layer, the data layout corresponding to each cloud storage system will be unevenly distributed, which will lead to a serious decline in the performance of data access [13–15]. After the storage system has been running for a period of time, it will become normal that the cor-responding data load is seriously unbalanced. Once the corresponding load of a data storage service provider is very high, the system may be blocked and stagnant. When the data are stagnant, the corresponding outside world must send out sudden data access and the whole system will have a bottleneck [16, 17]. Based on this, we aim to realize the load balancing of IoT terminal nodes to make the system achieve load balancing and realize the application opti-mization access function.

Based on the disadvantages of the corresponding data distribution and storage of IoT terminals considering the access cost, this article will propose a complex perceptual data placement algorithm based on the cloud storage dis-tribution of IoT terminal nodes and realize the I/O mode of accurately dividing IoT data by reasonably configuring hybrid perceptual algorithm and adaptive perceptual algo-rithm to optimize further the performance of the IoT data storage system considering the access cost. Aiming at the problem of terminal node load balancing, this article in-novatively proposes the terminal node data sorting and distribution algorithm, which realizes the accurate seg-mentation of the IoT data to be processed through the node data sorting and distribution algorithm to further accelerate the data reading and processing speed. At the same time, a periodic load balancing algorithm is added to the terminal node data sorting and distribution algorithm so as to further solve the problem of unbalanced node data layout. Based on the proposed algorithm, this article designs a load balancing cloud storage data distribution optimization system of IoT terminal nodes considering access cost and carries out ex-perimental verification in a real environment. The experi-mental results show that the data pattern division accuracy corresponding to the distribution strategy proposed in this article is improved to 97.13% and the corresponding data access efficiency is improved to 98.3% compared with the traditional distribution strategy. The data distribution strategy proposed in this article has obvious performance advantages and further promotion value.

The main structure of this article is as follows: the second section will mainly analyze the current research status of the load balancing cloud storage data distribution system of IoT terminal nodes considering access cost. The third section will focus on the analysis of complex perceptual data placement algorithm and terminal node data sorting and distribution algorithm and design the cloud storage data distribution optimization system for load balancing of IoT terminal nodes considering access cost. The fourth section will verify the system and analyze its data. Finally, the article is summarized.

2. Related Research: Research Status of Load Balancing Cloud Storage Data Distribution System of Internet of Things Terminal Nodes considering Access Cost

A large number of researchers and scientific research in-stitutions have studied and analyzed the data placement and load balancing problems faced by the load of terminal nodes of the IoT. At the research level of data placement and distribution, relevant cloud storage manufacturers in the United States put forward simple storage services, and their corresponding storage mainly involves three indicators: basic storage unit, unique identifier, and storage object container. When there are a few corresponding IoT data, the corresponding distribution strategy is relatively simple; therefore, not too much data will be stored in the corre-sponding storage object container. However, once there is

too much or blocked data, the performance of the cloud storage system will decline, which is not conducive to the maintenance and dynamic development of the system [18, 19]. European cloud storage service providers propose a static directory layering strategy, which mainly adopts a two-tier static object distribution model, alleviating the problem of too centralized data distribution on nodes to a certain extent, but this model still has certain limitations. The even distribution of the corresponding data depends on the corresponding IoT data level [20]. At the level of corresponding load balancing algorithms, relevant Chinese researchers have proposed a scheme to realize load balancing distribution in a distributed environment. However, this scheme does not consider the corresponding data copies, which may lead to multiple data copies appearing on the same data node after the relevant migration. At the same time, this load balancing strategy does not explicitly specify how much data can be migrated during the migration process. If the corresponding migration amount is too large, the load node will exceed the threshold [21]. Relevant European researchers have proposed a load balancing algorithm based on private cloud storage, which mainly realizes the balanced storage of data by calculating the weight corresponding to the storage node. However, the algorithm does not fully consider the heterogeneity of the node itself in the operation process. It treats all nodes equally, and it does not fully consider the heavy data at this time; if the data is too heavy, data migration will lead to the risk of increasing system bandwidth [2, 22]. Based on the analysis of the above research status, the current traditional load data placement algorithm and load balancing algorithm of IoT terminal nodes have more or fewer disadvantages, which leads to the accuracy of data mode division of cloud storage system and then leads to a serious decline in the corresponding data access efficiency.

3. Research on Load Balancing Cloud Storage Data Distribution Strategy of Internet of Things Terminal Nodes considering Access Cost

This section mainly analyzes and studies the corresponding key algorithms and system design of the IoT terminal node load balancing cloud storage data distribution system considering access cost. The corresponding system principle block diagram is shown in Figure 1. As shown in Figure 1, the key algorithms proposed in this article mainly have two points: the complex aware data placement algorithm and node load balancing strategy for processing data placement. At the corresponding hardware system design level, it also includes the distribution of IoT terminal nodes, IoT network element data management layer, IoT file service layer, and IoT data node control layer. There is no data management layer in the corresponding system, including IoT user and equipment management, IoT user data node allocation, IoT monitoring data management equipment, and IoT migration data management. The corresponding IoT file service layer includes the upload and download of IoT data files, the upload and download of quantity network audio and video files, the deletion of IoT files and other related modules. The whole system includes monitoring system components.

3.1. Complex Perceptual Data Placement Algorithm. The complex perceptual data placement algorithm proposed in this article mainly solves the problem of data placement distribution in the IoT. The complex perceptual data placement algorithm is mainly divided into two processing levels. Firstly, the I/O mode of the terminal data nodes of the IoT is identified and classified based on the decision tree algorithm. Then, the placement strategy is formulated based on the processed IoT data. The principle block diagram of the corresponding complex perceptual data placement algorithm is shown in Figure 2.

The IoT data I/O mode discrimination and classification algorithm based on decision tree mainly discriminates the data I/O type based on the six key factors of IoT data. The corresponding six factors are as follows: IoT data size, data type, IoT data life cycle, IoT data creation time, the last creation time, and the last modification time of IoT data [23–25]. Based on the above six factors, the I/O mode of IoT data is determined. The modes determined in this section are as follows: read-only mode, write-only mode, read more and write less mode, read less and write more mode. The IoT data I/O mode discrimination and classification algorithm based on decision tree mainly follows the following operation steps:

Step 1: calculate the gain value of IoT terminal node data based on six factors: IoT data size, data type, IoT data life cycle, IoT data creation time, IoT data last creation time, and last modification time. The corresponding calculation formulas are shown in formulas (1), (2), and (3), the corresponding IoT data size is V, the data type is L, the life cycle of IoT data is S, the creation time of IoT data is C, the last creation time of IoT data is L, and the last modification time is D.

$$\begin{aligned} G(\text{data}, D) = & Q(\text{data}_1 + \text{data}_2 + \cdots + \text{data}_i) \\ & - Q(\text{data}_1|D_1) + \cdots Q(\text{data}_i|D_i), \end{aligned} \tag{1}$$

$$\begin{aligned} Q(D) = & -[\eta_1(a_1)\log(\eta_1(a_1))) + \cdots \\ & + \eta_i(a_i)\log(\eta_i(a_i))), \end{aligned} \tag{2}$$

$$\begin{aligned} Q(\text{data}|D) = & \eta(D_1) * Q(\text{data}_1|D_1) + \cdots \\ & + \eta(D_i) * Q(\text{data}_i|D_i). \end{aligned} \tag{3}$$

Step 2: divide the I/O mode of the IoT data according to the gain value calculated in Step 1, and repeat the operations of Step 1 and Step 2 for the IoT data until all the IoT data can be defined as unique data mode.

Step 3: generate a data I/O mode classification model of IoT terminal nodes based on a decision tree to determine the complete data I/O type.
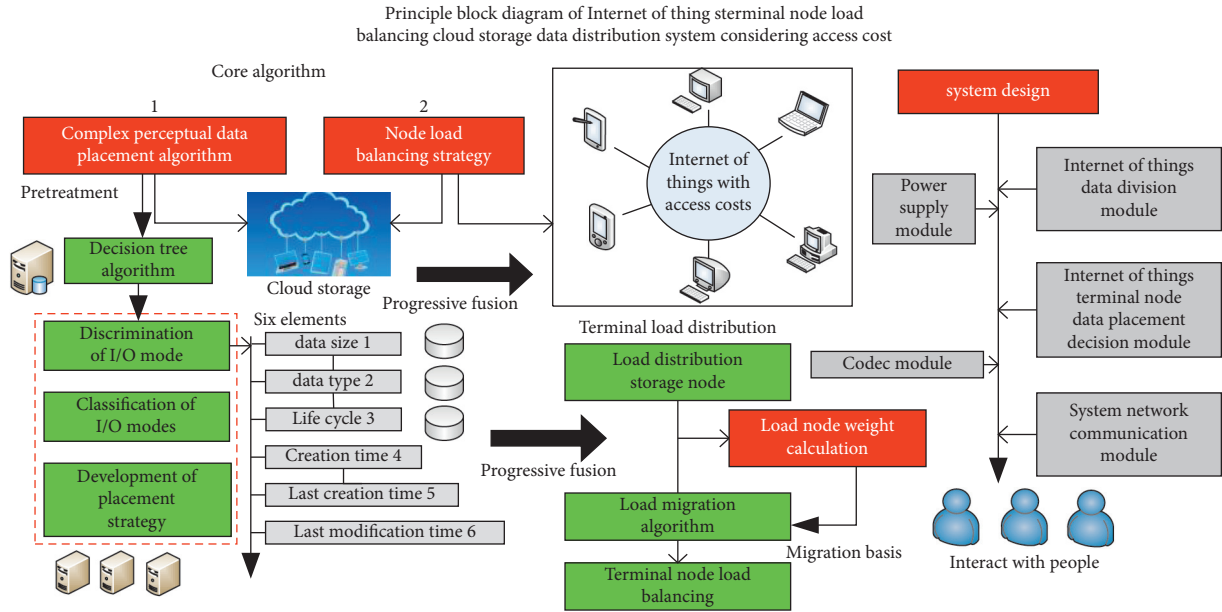
Figure 1: Principle block diagram of Internet of Things terminal node load balancing cloud storage data distribution system considering access cost.
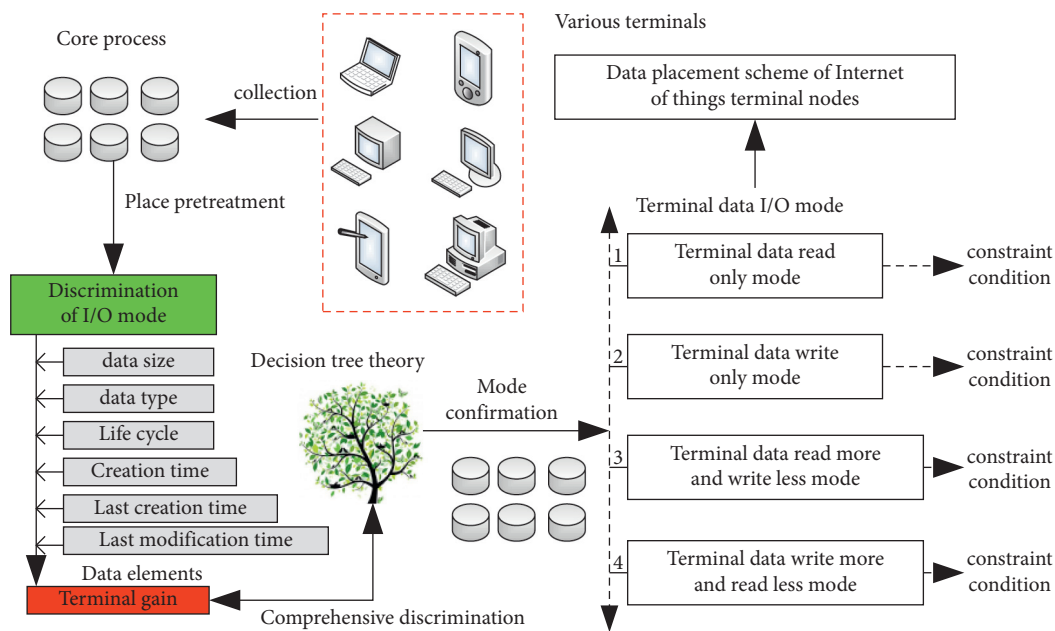
Figure 2: Principle block diagram of complex perceptual data placement algorithm.

Step 4: continuously collect the data to be placed, collect its corresponding six eigenvalues, call the corresponding decision tree IoT terminal node data I/O mode classification model, and constantly put new prediction results into the model so as to improve the accuracy of the model until all IoT terminal node data can be finalized.

After being processed by the I/O mode discrimination and classification algorithm of IoT data in the decision tree, it enters the link of the adaptive data placement algorithm. This algorithm is mainly based on the I/O mode type of IoT data to further determine its corresponding placement strategy and placement scheme. Based on this assumption, the read more write less mode is mode 1, the read less write

multimode is mode 2, the read-only mode is mode 3, and the write-only mode is mode 4. Take the read more write less mode and read less write more mode as case modes, and formulate their corresponding placement strategies.

The placement strategy objective function corresponding to the read more write less mode is shown in formula (4).

$$S = \min\left[\frac{(L * G_1(\text{mode1}))}{(A * d * w)}\right], \tag{4}$$

$$L * G_1(\text{mode1}) = \max\frac{\left(C_1^4 * \lg_1 + \cdots + C_i^4 * \lg_i\right)}{m_i, \quad 1 < i < n}, \tag{5}$$

$$A_1 = \left(\prod i \in N_i(1 - a_1) * \prod i \in N_i(a_1) + \cdots + \prod i \in N_i(1 - a_i) * \prod i \in N_i(a_i)\right), \tag{6}$$

$$A_1^s + A_1^g + A_1^T < a * A_1, \tag{7}$$

$$A_1^T = \left(a_1^4 * P_{\eta_1}^T * G_{d_1}\right) + \cdots + \left(a_1^4 * P_{\eta_i}^T * G_{d_i}\right). \tag{8}$$

The placement strategy objective function corresponding to the read less write multimode is shown in formula (9). This mode is mainly relatively sensitive to write delay. The

This mode has large downlink traffic and is sensitive to the read delay of the corresponding IoT data. The corresponding data placement constraints in this mode are shown in formulas (5)–(7) and (8). In the corresponding formula, A represents the cost proportion threshold and C represents the downstream data cost.

corresponding data placement constraints in this mode are shown in formulas (10)–(12) and (13).

$$S = \min\left[\frac{(L * G_2(\text{mode2}))}{(A * d * w)}\right], \tag{9}$$

$$L * G_2(\text{mode2}) = \max\frac{\left(C_1^2 * \lg_1 + \cdots + C_i^3 * \lg_i\right)}{m_i, \quad 1 < i < n}, \tag{10}$$

$$A_2 = \left(\prod i \in N_i(1 - a_1) * \prod i \in N_i(a_1) + \cdots + \prod i \in N_i(1 - a_i) * \prod i \in N_i(a_i)\right), \tag{11}$$

$$A_2^s + A_2^g + A_2^T < a * A_2, \tag{12}$$

$$A_2^T = \left(a_1^2 * P_{\eta_1}^T * G_{d_1}\right) + \cdots + \left(a_1^2 * P_{\eta_i}^T * G_{di}\right). \tag{13}$$

Through the confirmation of the above data placement scheme, the data sorting and distribution of IoT terminal nodes can be further confirmed so as to further optimize the data placement of the system.

In summary, based on the above algorithm, the heterogeneity of IoT terminal node data can be fully considered so as to realize the rationality of IOT terminal node data placement considering access cost and then prepare for load balancing.

3.2. Terminal Node Load Balancing Strategy. The terminal node load balancing strategy proposed in this section mainly includes load distribution, storage nodes, and load migration. The principle block diagram of the corresponding load balancing strategy is shown in Figure 3.

From Figure 3, we can see the load distribution, storage nodes, and load migration of the entire wireless sensor network.

In data node allocation, the weight of terminal node data needs to be calculated first. The corresponding calculation formula is shown in formula (14). The corresponding S represents the data state of the node, the corresponding a represents the heterogeneous weight of the data node, W_s represents the storage space used by the corresponding node data, and W_n represents the number of IoT nodes currently allocated. When the corresponding data weight is larger, the comprehensive available resources of the corresponding data node are smaller.

$$W = \frac{(S * W_s)}{(W_T * A) * W_n}. \tag{14}$$
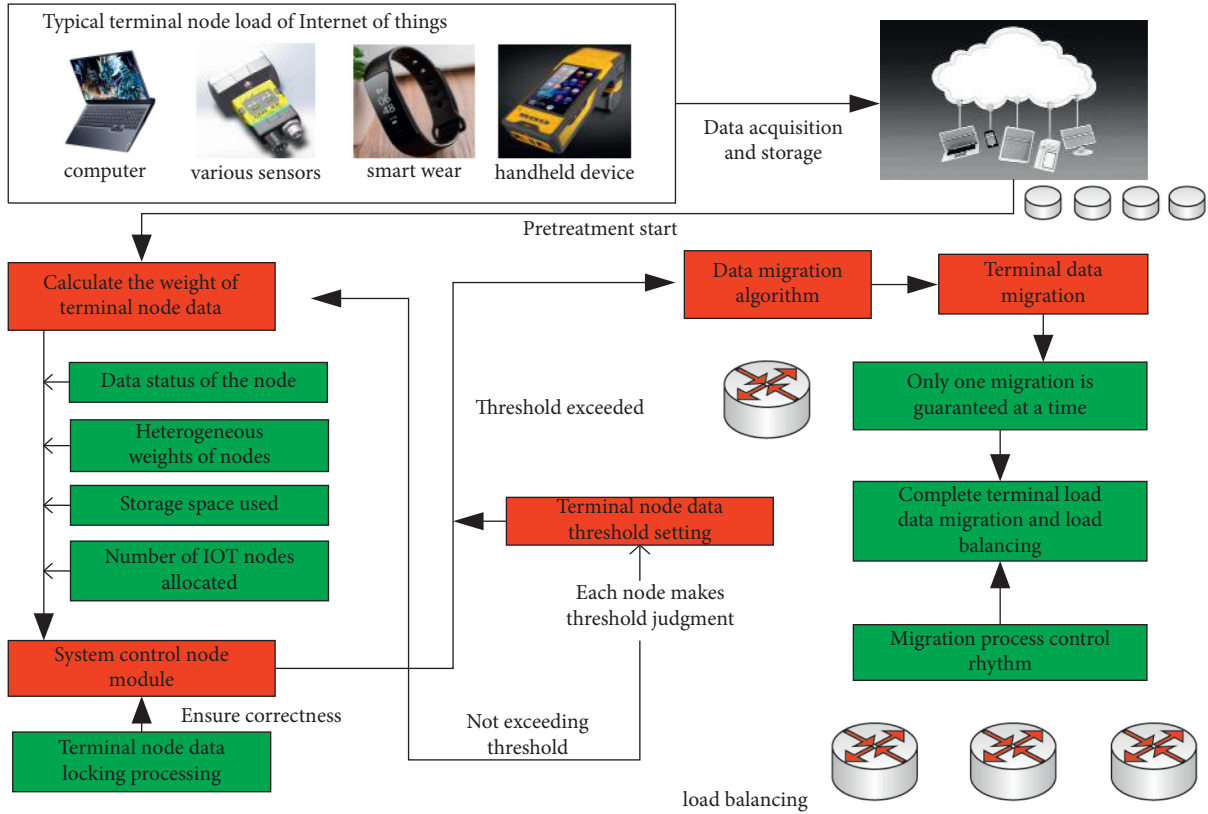
FIGURE 3: Principle block diagram of terminal node load balancing algorithm.

The corresponding data node allocation process is shown in Figure 4. It can be clearly seen from Figure 4 that each storage node corresponding to the wireless sensor network needs to continuously calculate the heterogeneous weight of its corresponding data node and feed the data back to the corresponding control node of the wireless sensor network in real-time. At the same time, the data also cover key information such as the storage space used by data nodes and the number of IoT nodes allocated. In order to ensure the correct rate of the algorithm of the entire system, the system is implemented with a locking scheme when multiple IoT data nodes are concurrent.

After completing the data node allocation, carry out the data migration operation. First, set the data load threshold as Wy, judge whether the load of IoT terminal nodes is too heavy based on this threshold, and carry out relevant migration actions. When the load weight of the corresponding cloud storage system is far less than the load threshold and the weight of the corresponding single load node is greater than the load weight of the whole system, notify the system to trigger the corresponding data migration algorithm for data migration. When the corresponding system load weight is much greater than the load threshold, notify the system to add the corresponding data storage node and notify the system that the load is too heavy at this time. In the actual data migration process, in order to ensure the correctness of data migration, the system needs to ensure that only one data migration operation is allowed at the same time, and the system needs to control the migration rhythm.

3.3. Design of Cloud Storage Data Distribution System for Load Balancing of Internet of Things Terminal Nodes Considering Access Cost.
The design framework of the cloud storage data distribution system for load balancing of IoT terminal nodes considering access cost is shown in Figure 5. It can be seen from the figure that the corresponding hardware system mainly includes IoT data division module, IoT terminal node data placement decision module, system network communication module, and IoT terminal node data encoding and decoding module.

At the design level of the corresponding IoT terminal node data division module, it mainly realizes the I/O mode division of the placed data. The module includes the data I/O model determination strategy, and the corresponding implementation method realizes the data type division model based on the decision tree through scikit-learn. The corresponding data placement decision module of the IoT terminal node is mainly a module to formulate the corresponding data placement strategy, which includes the data placement strategy algorithm. The corresponding network communication module is mainly to realize the interaction between multicloud storage, provide the function of cloud storage to obtain data and services from the upper server, provide the function of independent search and modification and other related operations of the IoT terminal, and provide the function of obtaining and monitoring cloud storage operators. The IoT terminal node data encoding and decoding module mainly helps restore relevant data. This module encodes the IoT
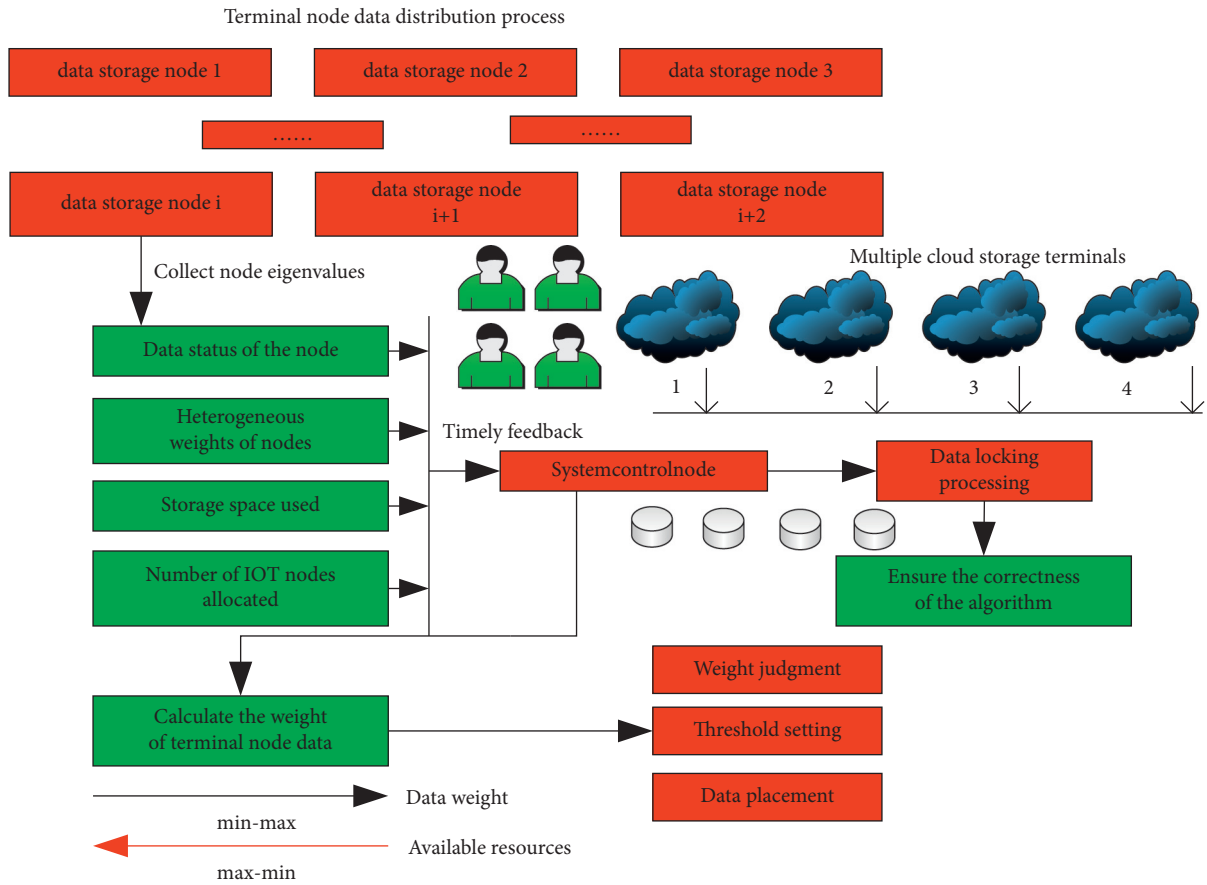
FIGURE 4: Principle block diagram of terminal node data distribution process.
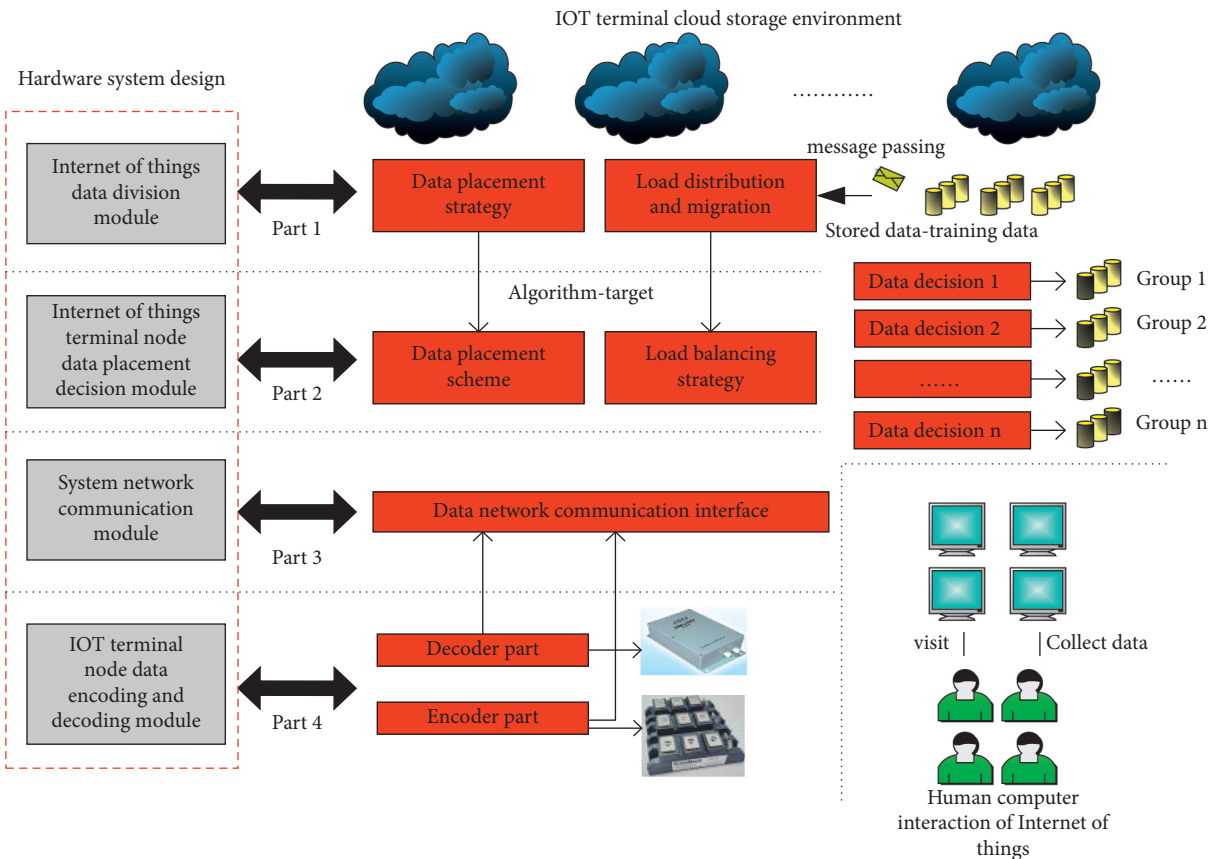
FIGURE 5: Design framework of cloud storage data distribution system for load balancing of Internet of Things terminal nodes considering access cost.
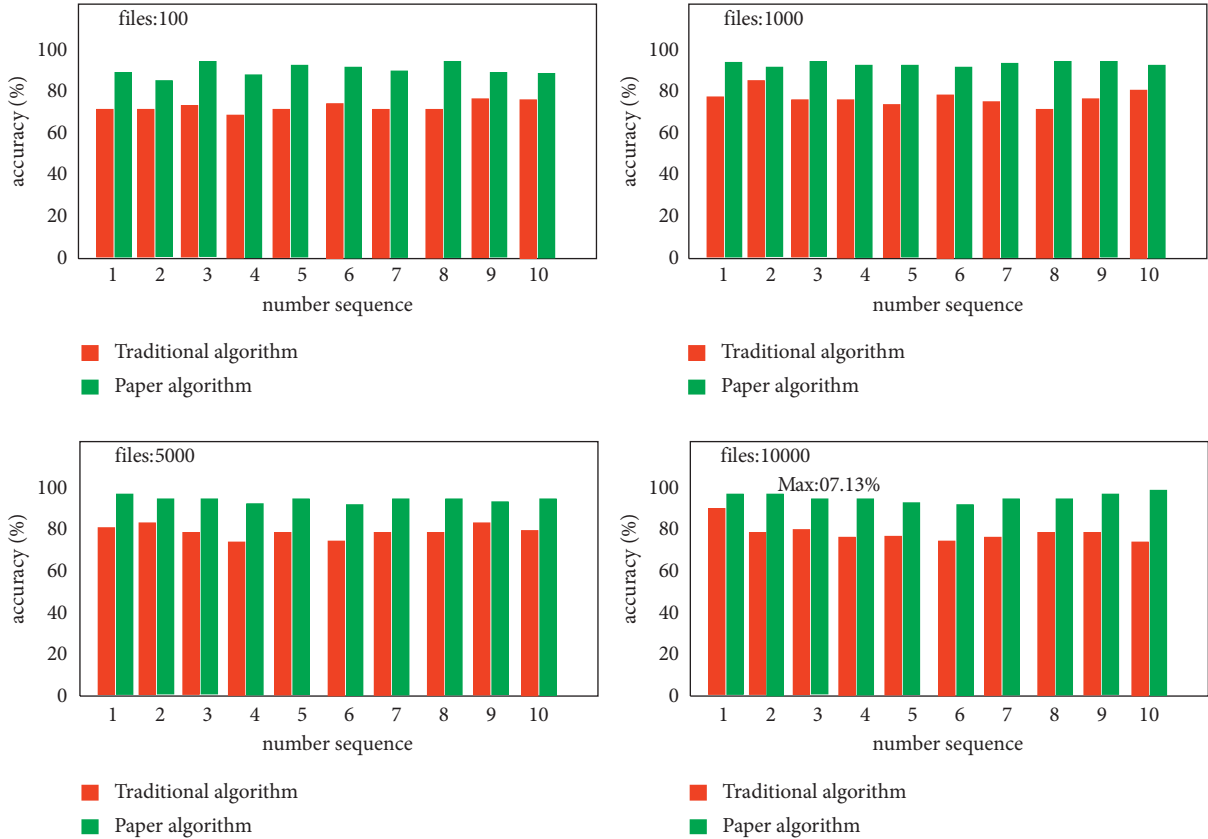
FIGURE 6: Column chart of data pattern division accuracy experiment.

terminal data through the encode() function and decodes the relevant IoT terminal data through the decode() function.

4. Experimental Verification and Data Analysis

In order to further verify the advantages of the actual system and the corresponding data distribution algorithm, this article is verified by the performance comparison experiment with the traditional distribution algorithm.

In the data mode division accuracy experiment, this article selects 100, 1000, 5000, and 10000 IoT terminal node files for comparative experiments. The corresponding experimental data are shown in Figure 6. It can be seen from the figure that the data distribution strategy proposed in this article is improving with the increasing number of files. When the number of files reaches 10000, the corresponding data partition accuracy is improved to 97.13%, which has obvious advantages over the traditional data distribution strategy.

Verify the response speed advantages of the distribution algorithm at the read, write, store, and delete levels when the number of files corresponds to 100, 1000, 5000, and 10000. The corresponding experimental results are shown in Figures 7(a), 7(b), and 7(c). As can be seen from the figure, at the corresponding write operation level, the write speed of the algorithm proposed in this article is about 15% higher than that of the traditional algorithm. At the corresponding read operation level, the write speed of the proposed

algorithm is about 20% higher than that of the traditional algorithm. At the corresponding storage operation level, the writing speed of the proposed algorithm is about 17% higher than that of the traditional algorithm. At the corresponding level of delete storage operation, the writing speed of the proposed algorithm is about 17% higher than that of the traditional algorithm. Based on the above data, it can be seen that the performance of the algorithm proposed in this article has been comprehensively improved compared with the traditional algorithm at the basic operational level.

In order to verify the advantages of this algorithm in the level of load balancing, the comparative verification is carried out when the number of files of IoT terminal nodes is 100, 1000, 5000, and 10000, and the corresponding load number of each storage node in the system is observed in the experiment. Under different file numbers, the corresponding load distribution is shown in Figure 8. It can be seen from the figure that the load distribution of the distribution algorithm proposed in this article is relatively uniform in each memory. At the same time, with the increase of the number of files, the number of files between the corresponding memories fluctuates little, while the corresponding load is seriously unbalanced under the corresponding traditional distribution strategy.

In order to verify the superiority of the algorithm proposed in this article in system data access efficiency, a comparative experiment is carried out based on the traditional distribution strategy. The experimental results are shown in Figure 9. It can be seen from the figure that the
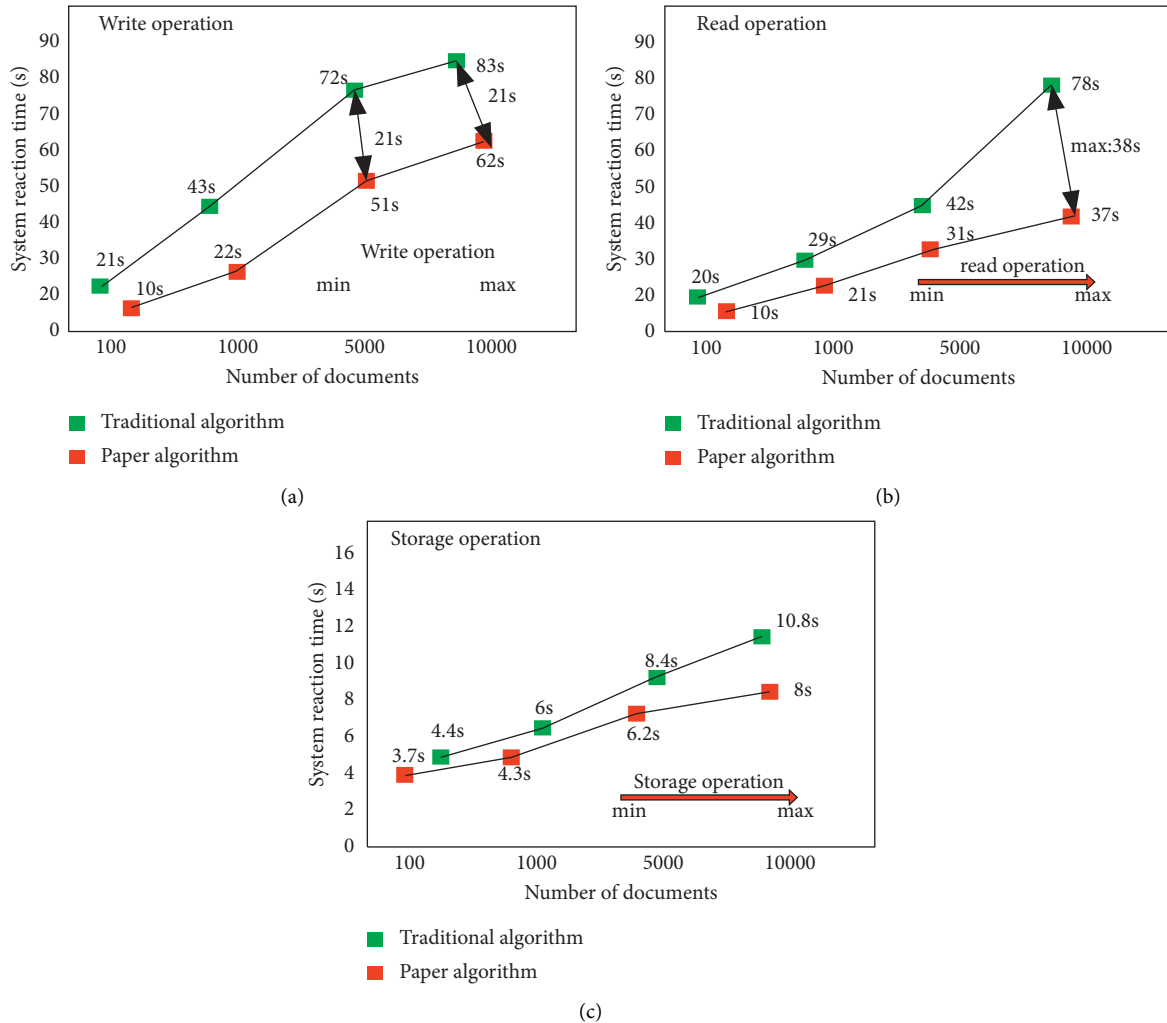
FIGURE 7: (a) Write operation level performance comparison column chart. (b) Read operation level performance comparison column chart. (c) Storage operation level performance comparison column chart.

distribution strategy proposed in this article has the highest system data access efficiency of 98.3%, which is nearly 20% higher than the traditional distribution algorithm. Therefore, this algorithm has obvious advantages over traditional algorithms.

The above analysis and experimental results further show that the algorithm proposed in this article has obvious advantages over the traditional algorithm in the cloud storage data system of load balancing of IoT terminal nodes considering access cost, and it can realize the optimization of data placement and load balancing. This greatly improves the accuracy of the system's data mode division of IoT terminal nodes and further improves the data access efficiency of the system.
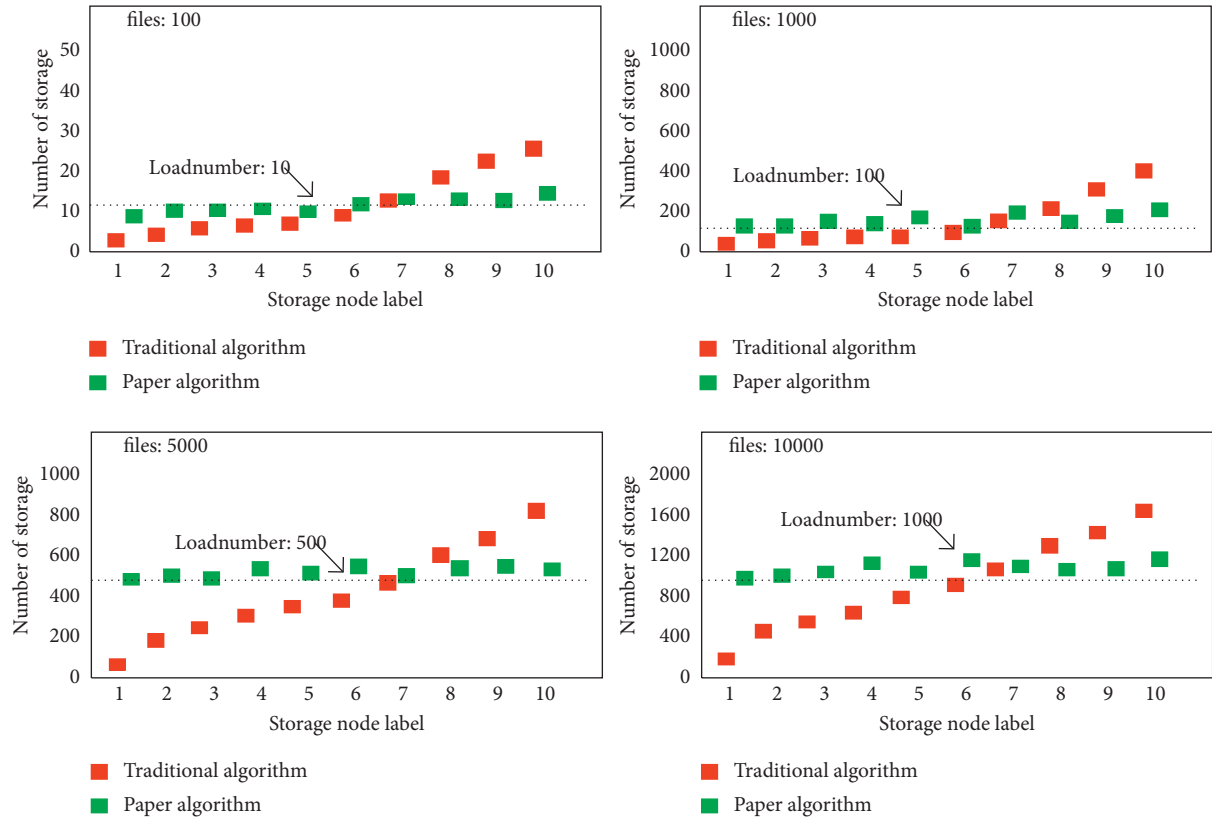
FIGURE 8: Broken line diagram of load distribution of each storage node under different numbers of files.
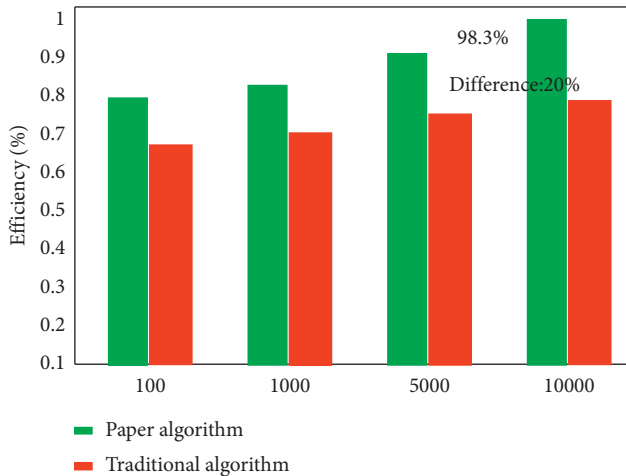
FIGURE 9: Comparison of system data access efficiency under different numbers of files.

5. Conclusion

This article mainly analyzes and studies the relevant research algorithms of cloud storage data distribution system with load balancing of IoT terminal nodes considering access cost and analyzes and compares their advantages and disadvantages. Based on the existing problems, this article proposes a complex sensing data placement algorithm based on the cloud storage distribution of IoT terminal nodes and realizes the accurate division of I/O mode of IoT data by reasonably configuring hybrid sensing algorithm and adaptive sensing algorithm so as to further optimize the performance of IoT data storage system considering access cost. Aiming at the problem of terminal node load balancing, this article innovatively proposes the terminal node data sorting and distribution algorithm. Through the node data sorting and distribution algorithm, the accurate segmentation of the IoT data to be processed is realized so as to further accelerate the data reading and processing speed. At the same time, a periodic load balancing algorithm is added to the terminal node data sorting and distribution algorithm. Thus, the problem of unbalanced node data layout is further solved. Based on the proposed algorithm, this article designs a load balancing cloud storage data distribution optimization system of IoT terminal nodes considering access cost and carries out experimental verification in a real environment. The experimental results show that the data pattern division accuracy corresponding to the distribution strategy proposed in this article is improved to 97.13% and the corresponding data access efficiency is improved to 98.3%, compared with the traditional distribution strategy. The data distribution strategy proposed in this article has obvious performance advantages and further promotion value. In the follow-up research, the article will focus on analyzing the energy consumption of wireless sensor networks. At the same time, it will further optimize the load balancing situation in the load balancing strategy. Moreover, it will further discuss the load balancing strategy problem of the traditional cloud storage network considering the cost.

Data Availability

The data used to support the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Quezada-Naquid, R. Marcelín-Jiménez, and J. L. Gonzalez-Compeán, "RS-Pooling: an adaptive data distribution strategy for fault-tolerant and large-scale storage systems," *The Journal of Supercomputing*, vol. 72, no. 2, pp. 1–21, 2015.

[2] M. Li, J. Zhang, J. Wan et al., "Distributed machine learning load balancing strategy in cloud computing services," *Wireless Networks*, vol. 26, no. 8, pp. 5517–5533, 2020.

[3] J. Huang, W. Liu, and Y. Su, "Load balancing strategy and its lookup-table enhancement in deterministic space delay/disruption tolerant networks," *Advances in Space Research*, vol. 61, no. 3, pp. 811–822, 2017.

[4] B. Cui, Z. Liu, and L. Wang, "Key-aggregate searchable encryption (KASE) for group data sharing via cloud storage," *IEEE Transactions on Computers*, vol. 65, no. 8, pp. 2374–2385, 2016.

[5] J. Kumar and A. Singh, "Secure and energy aware load balancing framework for cloud datacenter networks," *Electronics Letters*, vol. 55, no. 9, pp. 11–23, 2019.

[6] A. Ghasemi and A. T. Haghighat, "A multi-objective load balancing algorithm for virtual machine placement in cloud data centers based on machine learning," *Computing*, vol. 102, no. 9, pp. 2049–2072, 2016.

[7] A. Cucchetti, C. Ricci, G. Ercolani et al., "Efficacy and cost-effectiveness of immediate surgery versus a wait-and-see strategy for sporadic nonfunctioning T1 pancreatic endocrine neoplasms," *Neuroendocrinology*, vol. 101, no. 1, pp. 25–34, 2015.

[8] C. T. Yang, W. C. Shih, and C. L. Huang, "On construction of a distributed data storage system in cloud," *Computing*, vol. 98, no. 1, pp. 93–118, 2016.

[9] J. P. B. Mapetu, L. Kong, and Z. Chen, "A dynamic VM consolidation approach based on load balancing using Pearson correlation in cloud computing," *The Journal of Supercomputing*, vol. 77, no. 6, pp. 5840–5881, 2021.

[10] S. S. Sefati, M. Mousavinasab, and R. Z. Farkhady, "Load balancing in cloud computing environment using the Grey wolf optimization algorithm based on the reliability: performance evaluation," *The Journal of Supercomputing*, vol. 11, no. 3, pp. 1–25, 2021.

[11] J. Prassanna and N. Venkataraman, "Adaptive regressive holt–winters workload prediction and firefly optimized lottery scheduling for load balancing in cloud," *Wireless Networks*, vol. 34, no. 3, pp. 2313–2322, 2019.

[12] C. Liu, R. Ranjan, C. Yang, X. Zhang, L. Wang, and J. Chen, "MuR-DPA: top-down levelled multi-replica merkle hash tree based secure public auditing for dynamic big data storage on cloud," *IEEE Transactions on Computers*, vol. 64, no. 9, pp. 2609–2622, 2015.

[13] J. K. Liu, K. Liang, W. Susilo, J. Liu, and Y. Xiang, "Two-factor data security protection mechanism for cloud storage system," *IEEE Transactions on Computers*, vol. 65, no. 6, pp. 1992–2004, 2016.

[14] Z. Yin, X. Li, Z. Yao et al., "Divergences in reproductive strategy explain the distribution ranges of Vallisneria species in China," *Aquatic Botany*, vol. 132, no. 3, pp. 41–48, 2016.

[15] A. Hussain, M. Aleem, M. A. Iqbal, and M. A. Islam, "SLA-RALBA: cost-efficient and resource-aware load balancing algorithm for cloud computing," *The Journal of Supercomputing*, vol. 75, no. 10, pp. 6777–6803, 2019.

[16] S. Vazquez, P. Acuna, R. P. Aguilera, J. Pou, J. I. Leon, and L. G. Franquelo, "DC-link voltage-balancing strategy based on optimal switching sequence model predictive control for single-phase H-npc converters," *IEEE Transactions on Industrial Electronics*, vol. 4, no. 19, pp. 11–19, 2019.

[17] A. Tchernykh, M. Babenko, N. Chervyakov et al., "AC-RRNS: anti-collusion secured data sharing scheme for cloud storage," *International Journal of Approximate Reasoning*, vol. 102, no. 11, pp. 60–73, 2018.

[18] Y. Li and Y. Wang, "Synergistic strategy for the geographical traceability of wild Boletus tomentipes by means of data fusion analysis," *Microchemical Journal*, vol. 140, no. 34, pp. 38–46, 2018.

[19] B. Hwa, Z. Yuan, and G. Kce, "Functional broadcast encryption with applications to data sharing for cloud storage," *Information Sciences*, vol. 502, no. 78, pp. 109–124, 2019.

[20] A. A. Periola, O. A. Osanaiye, and A. T. Olusesi, "Future cloud: spherical processors for realizing low-cost upgrade in underwater data centers," *The Journal of Supercomputing*, vol. 32, no. 2, pp. 1–27, 2021.

[21] K. Karmakar, R. K. Das, and S. Khatua, "An ACO-based multi-objective optimization for cooperating VM placement in cloud data center," *The Journal of Supercomputing*, vol. 91, no. 9, pp. 1–29, 2021.

[22] Z. Zhou, P. Q. Tran, and K. Kieft, "Genome diversification in globally distributed novel marine Proteobacteria is linked to environmental adaptation," *The ISME Journal*, vol. 141, no. 8, pp. 1–18, 2020.

[23] J. Wu, L. Shen, and L. Liu, "LSH-based distributed similarity indexing with load balancing in high-dimensional space," *The Journal of Supercomputing*, vol. 76, no. 1, pp. 636–665, 2020.

[24] R. Regaieg, M. Koubàa, and Z. Ales, "Multi-objective optimization for VM placement in homogeneous and heterogeneous cloud service provider data centers," *Computing*, vol. 11, no. 3, pp. 1–25, 2021.

[25] M. Riahi and S. Krichen, "A multi-objective decision support framework for virtual machine placement in cloud data centers: a real case study," *The Journal of Supercomputing*, vol. 74, no. 7, pp. 2984–3015, 2018.

Hindawi

Research Article

Score Prediction of Sports Events Based on Parallel Self-Organizing Nonlinear Neural Network

Junyao Ling ⓘ

Xi'an University of Finance and Economics, Xi'an 710100, Shaanxi, China

Correspondence should be addressed to Junyao Ling; tyjxb@mail.xaufe.edu.cn

This paper introduces the basic concepts and main characteristics of parallel self-organizing networks and analyzes and predicts parallel self-organizing networks through neural networks and their hybrid models. First, we train and describe the law and development trend of the parallel self-organizing network through historical data of the parallel self-organizing network and then use the discovered law to predict the performance of the new data and compare it with its true value. Second, this paper takes the prediction and application of chaotic parallel self-organizing networks as the main research line and neural networks as the main research method. Based on the summary and analysis of traditional neural networks, it jumps out of inertial thinking and first proposes phase space. Reconstruction parameters and neural network structure parameters are unified and optimized, and then, the idea of dividing the phase space into multiple subspaces is proposed. The multi-neural network method is adopted to track and predict the local trajectory of the chaotic attractor in the subspace with high precision to improve overall forecasting performance. During the experiment, short-term and longer-term prediction experiments were performed on the chaotic parallel self-organizing network. The results show that not only the accuracy of the simulation results is greatly improved but also the prediction performance of the real data observed in reality is also greatly improved. When predicting the parallel self-organizing network, the minimum error of the self-organizing difference model is 0.3691, and the minimum error of the self-organizing autoregressive neural network is 0.008, and neural network minimum error is 0.0081. In the parallel self-organizing network prediction of sports event scores, the errors of the above models are 0.0174, 0.0081, 0.0135, and 0.0381, respectively.

1. Introduction

With the mature application of neural network technology and the rapid development of Internet, the research of data prediction has emerged. Data prediction often faces some noisy, messy, and nonlinear data. Neural networks have good robustness, [1]. This article briefly introduces the main clustering algorithms in data prediction, conducts an in-depth study, and optimizes the SOM learning algorithm, which improves the network iteration to a certain extent. The speed of training layer expansion self-organizing mapping network (GHSOM) was studied, the gray relational analysis was introduced into the GHSOM network, and the GRAGHSOM algorithm was proposed. The experimental results show that the GRAGHSOM algorithm reflects the importance of each component of the sample vector in the model in the process of high-dimensional data clustering and can perform clustering more accurately [2–5].

Because the fitting of a single neural network to a nonlinear dynamic system is generally computationally intensive, slow in convergence, time-consuming, and poor in generalization ability, a modular neural network chaotic parallel self-organizing network prediction method is proposed. That is, the phase space after the reconstruction of the chaotic parallel self-organizing network is divided into multiple subspaces, and an unbalanced subnet strategy is adopted. A neural network with a different structure is used in each subspace to calculate and track the local attractor trajectory, and the prediction result is finally passed into integrated module output [6]. This method not only makes use of the characteristics of self-organization and self-learning of neural networks but also greatly improves its

computational performance compared with a single neural network. It can track the attractor trajectory more accurately in the subspace, thereby improving the overall prediction. Clustering refers to the grouping of objects in the database into meaningful subsets, so that the members in a class are as similar as possible, and the differences in members among different classes are as large as possible. Clustering has unsupervised learning capabilities and is widely used in many fields [7–10].

This article first reviews the commonly used BP neural network and self-organizing neural network in neural network models. Then referring to the article, a new self-organizing difference theoretical model was established on the basis of self-organizing, and a better neural network structure was found through comparison. A better model was found by comparing the number of different classes and the prediction error of the hidden layers. The reference article combines the self-organizing neural network with the autoregressive model, improves the weight initialization and learning rate of the neural network, and finds a model with less error by comparing the number of different categories. In the article, the BP neural network is selected as the comparison model to evaluate the effect of the new model. The single-even network and multi-even network forecast models of the forecast system are studied. Matlab 6.5 programming is used to train the breakout forecast system based on the nonlinear neural network, using the simple, practical, and easy-to-program Visual C++ software for interface design, and a visual forecasting system based on nonlinear neural networks was developed. Experiments have proved that the use of a breakout prediction system based on a nonlinear neural network for breakout prediction can improve the prediction accuracy.

2. Related Work

Therefore, in recent years, the prediction method of chaotic parallel self-organizing network based on a neural network has obtained preliminary theoretical and application research results. However, on the one hand, most of the current researches are focused on the optimization of the network structure itself and the improvement of the two independent processes of phase-space reconstruction parameter selection, without considering the two as a whole. The reconstructed phase umbrella is divided into subspaces in order to track the attractor's trajectory with higher precision locally.

Torrents et al. [11] believe that the phase space of the system can be reconstructed from the parallel self-organizing network of a variable, and all the information of the entire dynamic system is contained in the chaotic parallel self-organizing network itself. The embedding theorem of Grossberg [12] further proves the idea and provides a theoretical basis for predicting the chaotic time series. However, how to construct the prediction model according to the phase reconstruction theory, select the reconstruction phase-space parameters, and predict model parameters is a key problem in the prediction of chaotic parallel self-organizing networks. For this reason, many nonlinear prediction methods for

predicting chaotic parallel self-organizing networks have been proposed. Among them, the neural network-based prediction method has distributed processing, self-organization, and self-adaptation. The excellent characteristics of self-learning columns and fault tolerance can better deal with uncertain and nonlinear prediction problems, so more research and applications have been obtained in the prediction of chaotic parallel self-organizing networks. Mohammadi et al. [13] used the neural network global method to predict the chaotic parallel self-organizing network. The results proved that when the MLP neural network is used as a global prediction method to predict the chaotic parallel self-organizing network, the effect is better than most local prediction methods. RBF networks have a strong biological background and are often used in the prediction of chaotic parallel self-organizing networks. Researchers are committed to improving the hidden layer node function and number selection of classic RBF networks.

The initial global prediction methods for chaotic parallel self-organizing networks mostly used feed-forward neural networks. There are many types of feed-forward networks, mainly including multilayer perceptrons (MLP) neural networks and radial basis functions neural network, nonlinear neural network (WNN), and fuzzy neural network (FNN). A representative MLP network is Park [14], which uses a multilayer feedforward network to simulate Lorenz chaotic sequences, where m represents phase-space reconstruction. The embedding dimension and its conclusion illustrate the important role of neural network in the prediction of chaotic time series. Yuan and Li [15] studied the gradient changes of the parallel self-organizing network. The experimental results show that the prediction performance of the classic RBF network has been improved [16]. The researchers proposed a cross-validation subspace method to select the optimal number of hidden layer nodes. This method is used in the prediction of noise-containing chaotic parallel self-organizing networks. The simulation object is the chaotic sequence of radar sea echoes. Nonlinear neural network is a new type of feedforward neural network based on nonlinear theory. Nonlinear neural network is compatible with the superiority of nonlinear and neural networks [17–20].

3. Construction of Sports Event Score Prediction Model Based on a Parallel Self-Organizing Nonlinear Neural Network

3.1. Parallel Self-Organizing Architecture Nesting. According to the different objects to be studied, the behaviors of parallel self-organizing networks are very different. Even for the same object, there will be different performances at different times. Therefore, in order to treat different time series in different ways, parallel self-organizing networks can be classified according to the way of data selection and the form of data changes over time. Figure 1 shows the nested parallel self-organizing architecture.

When using the combination of difference weighted average to predict the parallel self-organizing network, the
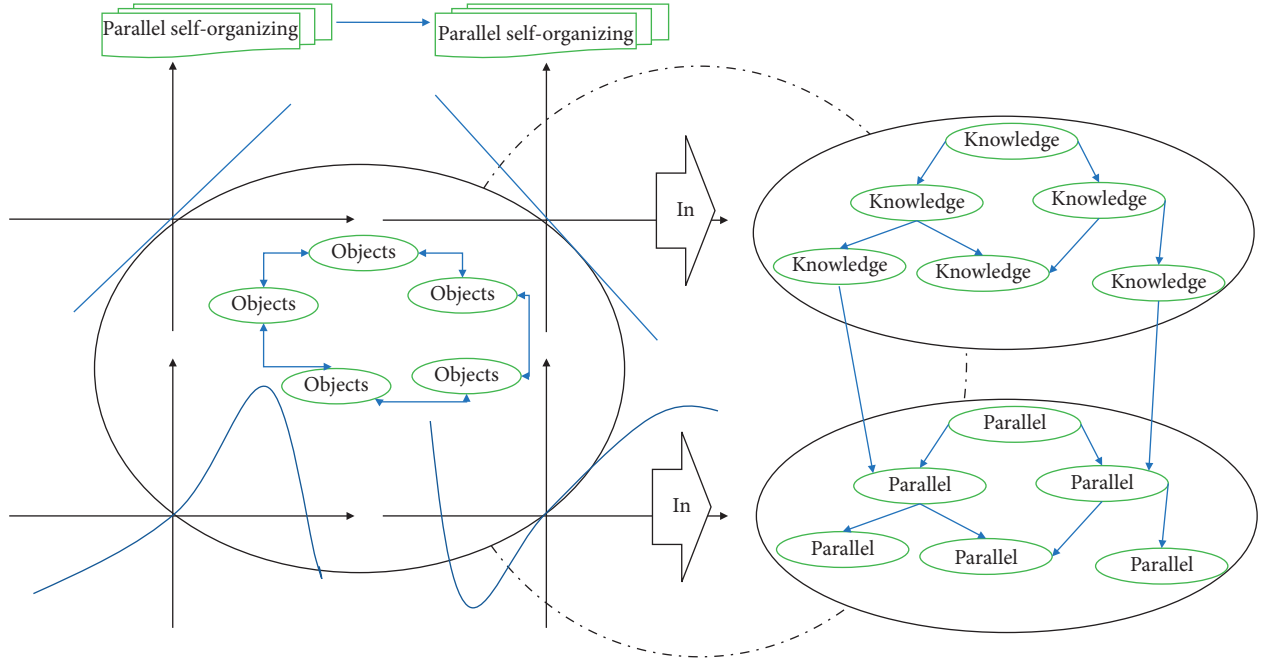
FIGURE 1: Parallel self-organizing architecture nesting.

self-organizing neural network to the historical data is first used and the input data are divided into several categories. The target value is calculated according to the formula to calculate the weighted average of the difference. When predicting the behavior of new data, first classify it according to the trained self-organizing neural network and then add the weighted average corresponding to the class of the data to the last value of the new data, which is the prediction corresponding to the new data.

$$R - \int_0^\infty \left(R_{FR} + R_{F1} - \frac{R_{FR} + R_{F1}}{R_{FR} R_{F1}} \right) dt = 0, \tag{1}$$

$$\exp\left[i \, df(w_i) \right] - \exp\frac{|D|}{\left| \{D_i: D_j \in w_j\} - 1 \right|} = 0. \tag{2}$$

First, we must establish a self-organizing neural network to classify the known data. The known data exist in the form of vectors. The elements of each vector are the parallel self-organizing network values selected from the parallel self-organizing network at a certain time interval. And all vector elements are selected in the same way. Let the dimension of the vector be l. The network weights of the initial self-organizing neural network are randomly selected. The risk value of the network is obtained by adding the risk value of the host

$$u(k, t, x(k, t)) = \begin{cases} \varepsilon(i, n) \times w(k, t), & 0 < t < k, \\ \delta(i, n) \times w(k, t), & t > k, \\ y(n) \times w(k, t), & t < 0, \end{cases} \tag{3}$$

$$\sum \nabla^2 \frac{\alpha - 1}{a^2} * \frac{\partial^2 \alpha}{\partial t^2} + \int \frac{\rho}{\varepsilon} dx dy = 0. \tag{4}$$

When all the known data are input to the self-organizing neural network, the optimal neuron is obtained, and the weight vector of the optimal neuron and its surrounding neurons are adjusted, and the self-organizing neural network completes one step of training. Then, the data are re-input into the self-organizing neural network for the next step of training, until the optimal neuron corresponding to each input vector in the two trainings before and after does not change, and the self-organizing neural network completes the training. Taking the weight vector of the neurons in the competitive layer after training as the clustering center, the data can be classified finally.

$$f(x_i, x_2, x_3) = G(F(x_1, W(x))) G(F(x_2, W(x))) G(F(x_3, W(x))). \tag{5}$$

The neurons receive the data passed by the neurons in the previous layer and adjust the data according to their own transformation functions. We perform the transformation and then pass the transformed data to the next layer of neurons. The last part receives the data of the last layer of hidden layer neurons and calculates the final result and error of this input.

3.2. Nonlinear Neural Network Node Output.

Structurally, the characteristics of nonlinear neural networks are as follows: (1) distributed storage: that is, all neurons of the neural network jointly complete the information processing function, and each neuron has an independent function to process data; (2) parallelism: that is, the ability of all neurons of the neural network to complete information processing concurrently; and (3) fault tolerance: a single neuron's error in information processing will not affect the entire system. In terms of ability, the neural network can spontaneously react to external input and output data and makes appropriate adjustments to its own network parameters according to the

input and output samples to adapt to the new input samples. Each neuron in the input layer stores an element of the input vector, and all neurons assign data values to the discriminant function of the neurons in the competition layer through the connection with the data according to the discriminant function, select the best matching neuron by comparing the value of the discriminant function of each neuron, and modify the network connection weight parameter according to the input data. The connections represent a certain distance between neurons in the competition layer. The distribution of this weight in one or two dimensions is fixed and will not change during training. Figure 2 shows the distribution of nodes in a nonlinear neural network.

In a forward network without feedback, once a signal passes through a neuron, the neuron's processing process ends. In an interconnected network, signals have to be repeatedly transmitted between neurons, and the network is in a dynamic state of constant change. Starting from a certain initial state, after several changes, it will reach a certain equilibrium state. According to the structure of the network and the characteristics of neurons, the operation of the network may also enter periodic oscillations or other equilibrium states such as chaos. This parallel self-organizing network can also be called the aggregate index parallel self-organizing network, which is a statistical sequence arranged in time, and the data content is the statistical value of the same phenomenon or characteristic value. From the absolute parallel self-organizing network, the overall level of the phenomenon to be studied in a certain period can be observed, and it can also reflect the level of the phenomenon at a certain point in time. Therefore, the absolute parallel self-organizing network can be further divided into time series and time-point series according to the time selection method. The changes in the value of the curvilinear parallel self-organizing network are also inclined, but the tendency is not fixed, the magnitude and direction of the change may change, and the change of the value is not periodic. The linear parallel self-organizing network value changes have a certain tendency; that is, within a long time range, the value of the parallel self-organizing network gradually rises or decreases, and an upward or downward trend can be observed.

3.3. Quantification of Score Prediction in Sports Events.
The normalization of artificial neural network sports score prediction requires proper processing of data to meet the special requirements of artificial neural network for data. According to different neuron transfer functions, many artificial neural network models only accept numerical data in the range of [0, 1] and [−1, 1]. Therefore, the data must be scaled down to this interval. Scalar numerical data are generally evenly distributed in a certain range and can be directly mapped to the interval [0, 1]. If the distribution of numerical data is not uniform, you can use piecewise linear equation or logarithmic equation to transform the data, and then scale down d, N to specify the interval. Discrete data are represented by encoding them with 0 and 1, or assigning a value to them in a specified continuous interval. Numerical

data represented by vectors or arrays can sometimes be processed in groups; that is, the vector as a whole is subjected to regularization processing. There are several regularization methods, the most commonly used is to calculate the square root of the sum of the squares of the elements, and then it is used to remove each element. The second method is to first find the sum of all the elements, and then it is used to divide each number. In this case, the sum of the elements after regularization is 1.0, and the value of each element represents their contribution to the group. The third method is to remove each element with the maximum value in the vector. Figure 3 shows the quantitative processing of score prediction for sports events.

Therefore, all of the above factors can be converted into time factors, and the parallel self-organizing network analysis method can be used to mathematically model the time factor, that is, the value of the score of a sports event. Through the historical value of the score of a sports event, it is possible to discover the change of the score of a sports event over time and predict the future development of sports event scores. On the other hand, the score of sports events is a relatively open data, which is relatively easy to obtain, which also provides convenience for research. The parallel self-organizing network of sports scores is a parallel self-organizing network with the value of sports scores as data elements. The scores of sports events can be taken every hour, every day, or other time as the parallel self-organizing network of sports scores. In competition, a vector is used as the input of the neural network. Each unit of the competition layer will form a discriminant function according to the weight of its own network and calculates the discriminant function with the input vector as the value of the independent variable, and the function value is the best. The neuron becomes the best matching neuron in this input. Other neurons surrounding the best matching neuron within a certain range will participate in the cooperation and become the neuron to be adjusted in the process of self-adaptation. Adaptive neurons within the topological neighborhood of the optimal neuron must modify their own network connection weights according to the value of the input vector to achieve the discriminative mode of these neurons.

3.4. Recursive Prediction Model Weight Factor.
Data preprocessing is an important step in the data prediction process. The quality of the data directly affects the quality of the mining effect. Data preprocessing includes data cleaning, data integration, data conversion, and data reduction. The task of data cleaning is to eliminate noise or inconsistent data. The inconsistency of the data will reduce the credibility of the data prediction results.

The risk value of the host is obtained by adding the risk value caused by the vulnerabilities of the related state nodes. The database construction module is responsible for processing the system configuration information and vulnerability information obtained by the scanning module, extracting their nonquantitative attributes and quantifiable attributes, forming an information database and a
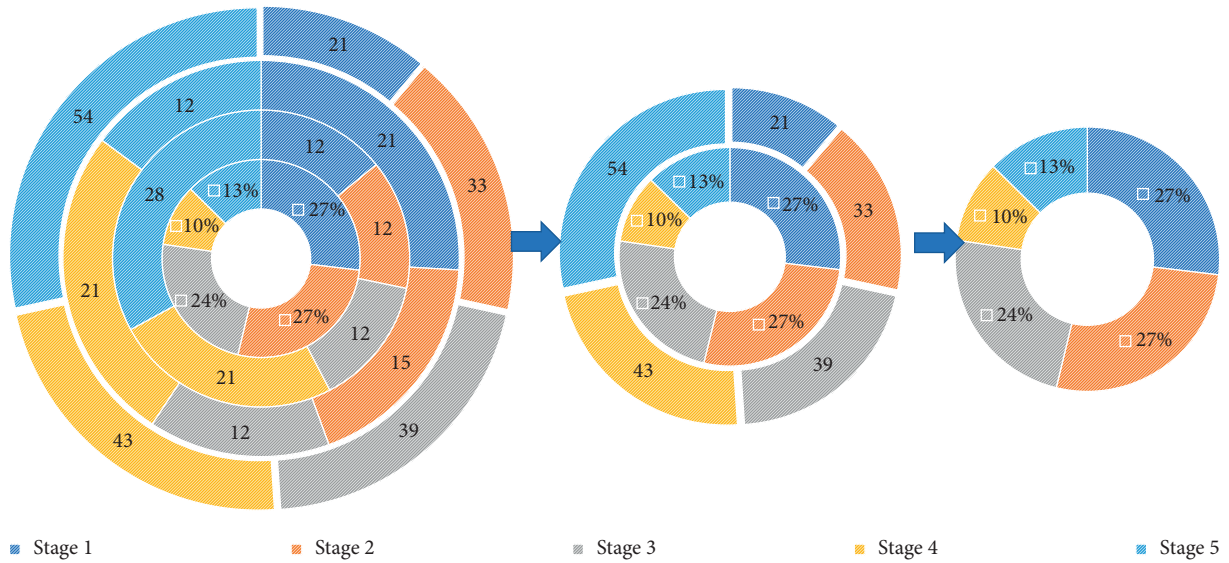
Figure 2: Node distribution of the nonlinear neural network.
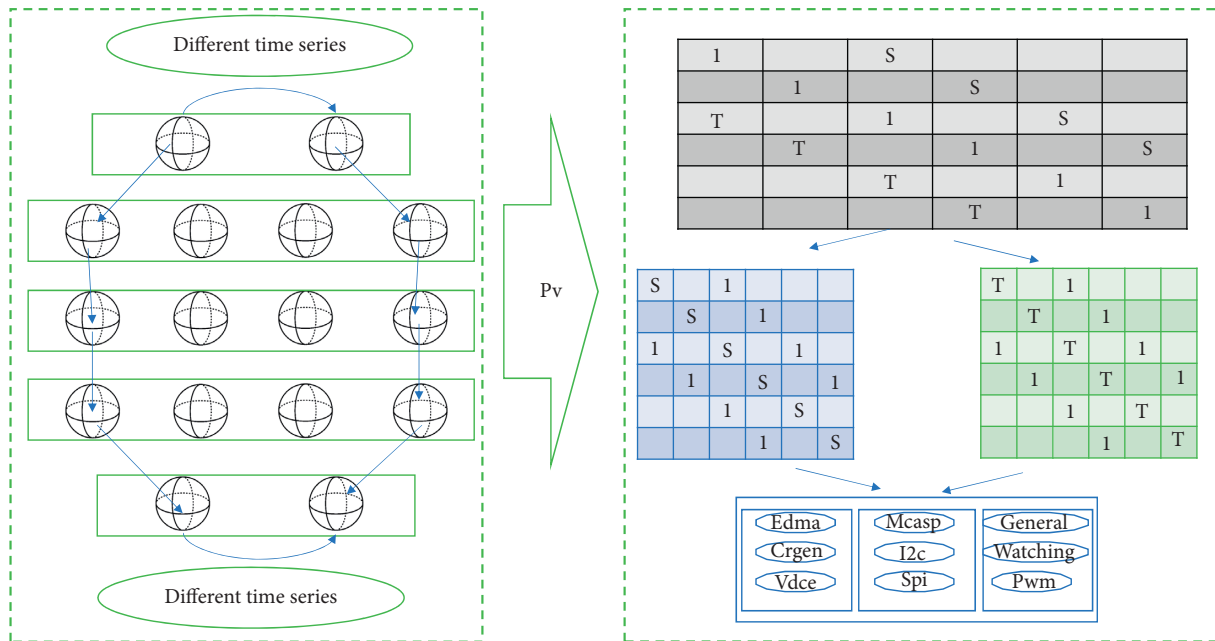
Figure 3: Quantitative processing of score prediction for sports events.

vulnerability database; it is also responsible for analyzing the relationship among vulnerabilities. Mining was performed to form a vulnerability database. The data used for data prediction may come from multiple actual systems, so there is a problem of heterogeneous data conversion. Because there may be many inconsistencies between the data of multiple data sources, data integration is not a simple copying process. Data transformation is the transformation or unification of data into a form suitable for mining, including data generalization, normalization, attribute construction, and smoothing. Data reduction includes operations such as data aggregation, attribute selection, and data compression. Figure 4 shows the clustering of the weight factors of the prediction model.

Multilayer perceptron is a popularization form of the single-layer perceptron. The information of the multilayer perceptron network is forwarded layer by layer, and each unit of the lower layer is connected to each unit of the upper layer. The input unit operates layer by layer according to the input/output relationship, and the connection weight between each layer can be adjusted by learning rules. It can be seen that the multilayer perceptron is actually a design of multiple single-layer perceptrons through appropriate combination, which can realize any shape division. In this method, each discrete class is assigned a value from 1 to N represented by a binary number. For example, if a variable has 32 possible values, it can be represented by a binary vector of length 5. As long as the discrete value is arbitrary
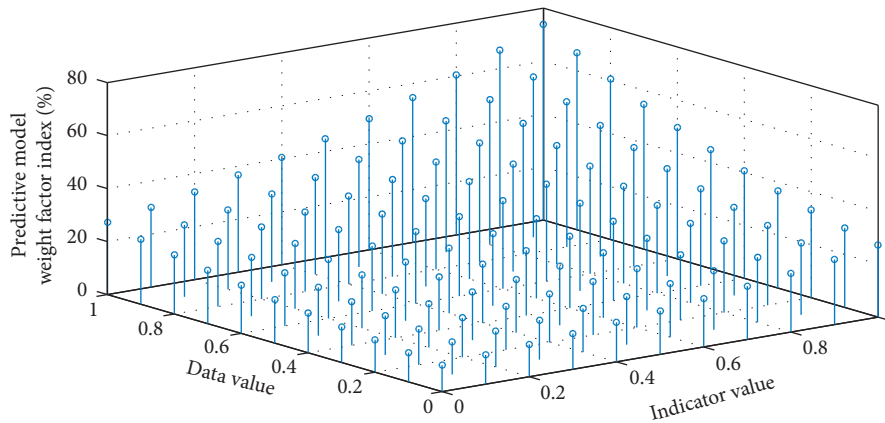
FIGURE 4: Weight factor clustering of the forecasting model.

and does not have any order, binary code is a better representation. However, when the discrete value is converted to a binary code, the bit value varies greatly. For example, the binary code of the seventh element in a discrete class is 000111, and the binary code of the eighth element is 001000. Hamming distance is a measure of the similarity between two binary numbers, expressed by the number of bits of different symbols in two codes. In this case, the Hamming distance between 7 and 8 is 4. If you want the neural network to think that the two input modes of 7 and 8 are similar, you must choose the thermometer code.

4. Application and Analysis of Sports Event Score Prediction Model Based on Parallel Self-Organizing Nonlinear Neural Network

4.1. Parallel Self-Organizing Nonlinear Neural Network Data Preprocessing. The nonlinear neural network selected in this paper has a hidden layer, and every 10 pieces of data are used as a set of inputs. In the training phase, a matrix composed of 500 groups of 10-dimensional input vectors is selected as the input matrix, and the matrix composed of the target value corresponding to each group of input vectors is used as the output matrix. The parameters are adjusted until the error reaches the target value of training times, stop training. In the prediction stage, the new data are used as the input vector, the corresponding predicted target value is calculated according to the adjusted BP neural network, and the actual target is compared to evaluate the performance in Figure 5.

For the measured parallel network, we calculate the ratio of pseudo-neighbors from the minimum embedding dimension. When the ratio of embedding dimension m to pseudo-neighbors is less than 5%, pseudo-neighbors no longer increase with embedding dimension. When it decreases, it can be considered that the singular attractor is fully expanded, and m at this time is the minimum embedding dimension. Generally speaking, when the distribution of phase-space points is relatively sparse, it is better to use the pseudo-neighbor point method, because this method can automatically adjust the distance scale with the sparse changes of the phase-space points so that the statistical results are more reliable. When the phase-space points are

densely distributed, it is more reasonable to use the saturated correlation dimension method.

To quantify the risk value of the state node, it is necessary to calculate the probability that the attacker can successfully attack to reach each state node. It can be seen that the average relative error of the PCA-GABP prediction model is both the smallest. The number less than 3% is 39, which is 2 more than the number where the absolute value of the relative error of the GABP model is less than 3%. After comparing and analyzing the data in the article, it can be concluded that the model whose input data have been processed by PCA has a higher prediction accuracy than the unprocessed model; at the same time, regardless of whether the input data has been processed by PCA, the BP neural network improved by GA prediction accuracy is significantly higher.

4.2. Realization of the Simulation of the Score Prediction Model for Sports Events. Also taking the sports event factor as an example, the GA nonlinear neural network has 48 inputs; that is, there are 48 neurons in the input layer of the network, and the output is also 48; that is, there are 48 neurons in the output layer of the network. Using the floating-point coding scheme, each connection weight or threshold is directly represented by a real number, and each individual contains a network ownership value and threshold. The weight of a network is composed of two parts.

Finally, the asset damage value caused by the related vulnerabilities is multiplied by the asset value of the host node to quantify the risk. This article has 48 input neurons and 48 output neurons. Assuming there are 36 hidden layers, plus the thresholds of the hidden layer and the output layer, the code length of each individual is $48 \times 36 + 36 \times 48 + 36 + 48 = 3540$. We input the training samples in the article into the prediction model, and after many debugging, it is finally determined that there are 36 nodes in the hidden layer. The simulation results of the GABP neural network model are as described in the text. Figure 6 shows the score forecasting training situation of sports events.

If the total input of the neuron is far away from the threshold, the actual output of the neuron is either the
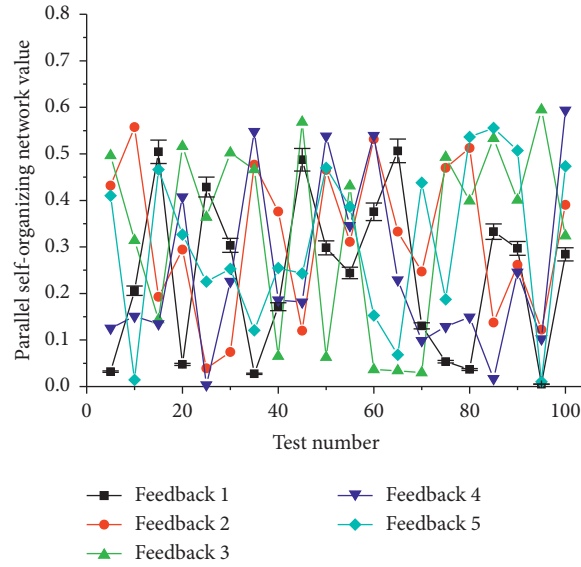
FIGURE 5: Distribution of connection points of parallel self-organizing nonlinear neural network.
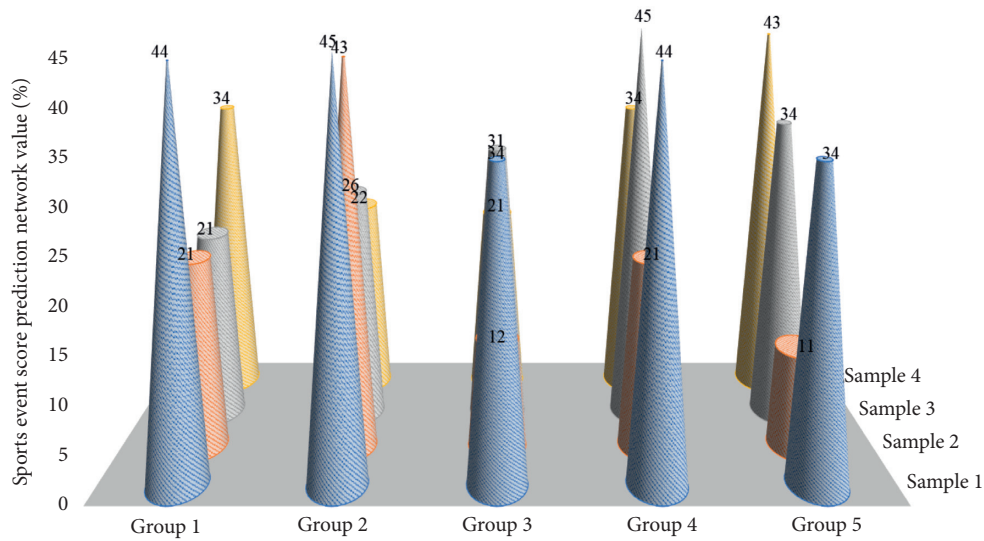
FIGURE 6: Score prediction and training situation of sports events.

maximum value or a small value due to the saturated nonlinear characteristics of the neuron. When the total input of a neuron is too far above the threshold, it is said to enter the saturation zone. When the total input falls into the saturation zone and the actual output contradicts the target output, it is incorrect saturation, that is, false saturation. At this time, there needs to be a large amount of modification to the weight, but in fact, because the derivative value is close to zero at this time. The amount of modification of the de-rivative weight is very small. The same is true for hidden layer neurons, but the error is calculated by error back propagation. Once learning enters the false saturation state, it will take a long time to get rid of this state or even unable to find the global optimum. Table 1 is the description of hidden layer neuron training.

Calculated by the pseudo-neighbor method, the mini-mum embedding dimension m is 9, the DRNN model is

TABLE 1: Description of hidden layer neuron training.

Index number	First level	Second level	Third level
		0.31	0.64
1	DRNN layer capability	0.22	0.21
		0.54	0.08
		0.24	0.13
2	DRNN layer ability	0.46	0.56
		0.11	0.34
3	Hidden layer neuron training	0.09	0.24
		0.27	0.32

constructed, the recursive delay is 1, and the same $9 \times 21 \times 1$ nonlinear neural network is constructed. The first 2000 points among the 3000 points are used as training data, and the last 1000 points are used as prediction data. Since the predicted curve is relatively close to the actual curve, it is

difficult to distinguish it from the graph, so only the absolute error $P(1)$ of the DRNN model and the nonlinear neural network prediction are given.

When network administrators find unknown vulnerabilities, they need to integrate the vulnerability information and submit it to the national information security vulnerabilities by e-mail. In the training phase, the network is used to layer the input data into several categories and calculate the weighted average difference corresponding to each category of data. In the prediction stage, the trained neural network is used to divide each group of input data into a certain classification area, and then, the prediction is made based on the calculated weighted average difference. The article lists the results of the self-organizing difference weighted average method to predict the results of the parallel self-organizing network of the sports event score, that is, when the number of historical data classifications is different.

4.3. Example Application and Analysis.

In the experiment, we set the output as $y(t+1)$ and then input as $x(t-4)$. The network has 5 nodes, and the time interval between each neuron is 1. In the experiment, 2500 data are generated on each chaotic parallel self-organizing network of the benchmark, and the whole of them is divided into three parts: training data, verification data, and test data. The ratio of the amount of data of the three parts is $14:6:5$, that is, 1,400 training data, 600 verification data, and 500 prediction data. It can be seen that the hidden layer nodes should be twice the input layer, that is, 10 nodes.

Then, the staff will analyze the vulnerabilities, and only after several working days can the information such as the category and hazard value of the vulnerabilities, be published on the official vulnerability platform. The output layer has only one neuron, that is, the predicted value of the stock closing index on the next day. The structure of the recurrent neural network is 51. The DRNN model in this paper is greatly improved relative to the nonlinear neural network model, and it can be seen that the fitting degree of the true exponential curve is also better than the fitting degree of the nonlinear neural network model. It can be seen that the error of the recursive model in this article is concentrated between $(-20, 20)$, the error value is small, the error of the nonlinear neural network model is between $(-80, 40)$, the error value is large, and some points deviate greatly from the true value, which is caused by the high degree of the feedforward network. Figure 7 shows the output effect of the sports event score prediction network.

This will not only cause a huge workload for the vulnerability management staff, but, more importantly, the network administrators will not be able to obtain information about unknown vulnerabilities in a timely manner. Because these feature variables may have high correlations, these features are redundant features. The results show that although the overall detection rate is very high, the U2R class is almost undetectable, and the R2L detection rate is relatively low. There is no doubt that this is because the number of samples is too small, which makes it difficult to build a
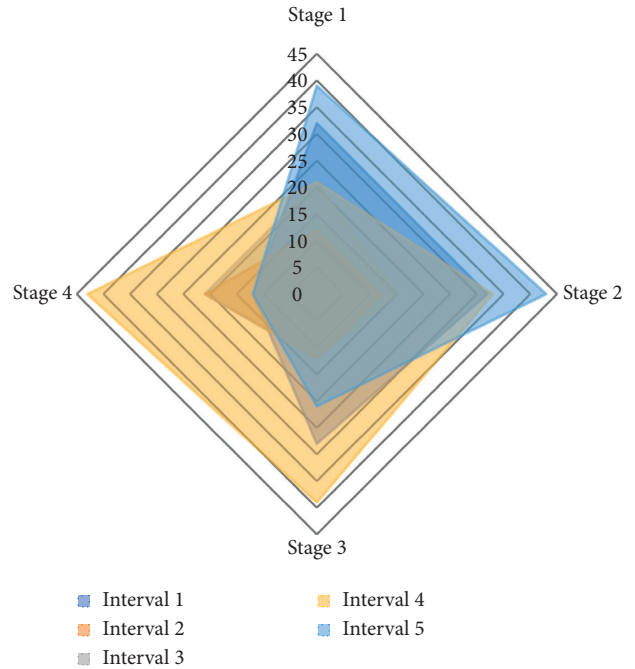
FIGURE 7: The output effect of the sports event score prediction network.

classifier for these attack classes. The error detection rate is high, and the two types are mostly detected as normal types, which are absolutely not allowed in IDS. To avoid this, deduplication is performed before training. In order to improve the performance of the classifier without greatly changing the number of samples in the data set, this article increases the number of U2R samples in the data set to ensure the basic classification performance. The system structure adopts the client/server structure. The server is responsible for scanning the hosts and devices in the network to obtain the configuration information and vulnerability information of all the hosts in the network; the client is responsible for collating the scanning results to form various databases. On this basis, certain mathematical methods are used to measure the overall risk assessment value of the network and determine the risk level. Therefore, in addition to filtering irrelevant features, the optimal feature subset must also filter redundant features. Figure 8 shows the convergence of the sports event score prediction network.

There are N individual individuals to form the initial population. The value of the population number N will affect the efficiency of the entire algorithm. If the value of N is too large, the calculation amount will increase, and the running speed will be too slow. If the value of N is too small, although the operating speed can be increased, the diversity of the population is reduced to a certain extent, which may cause premature convergence, that is, premature maturity. According to experience, the value of N is generally between 10 and 200.

Feature selection methods can affect machine learning performance, model complexity, and model running speed. One of the more popular feature selection methods is to select n variables that are highly correlated with categorical
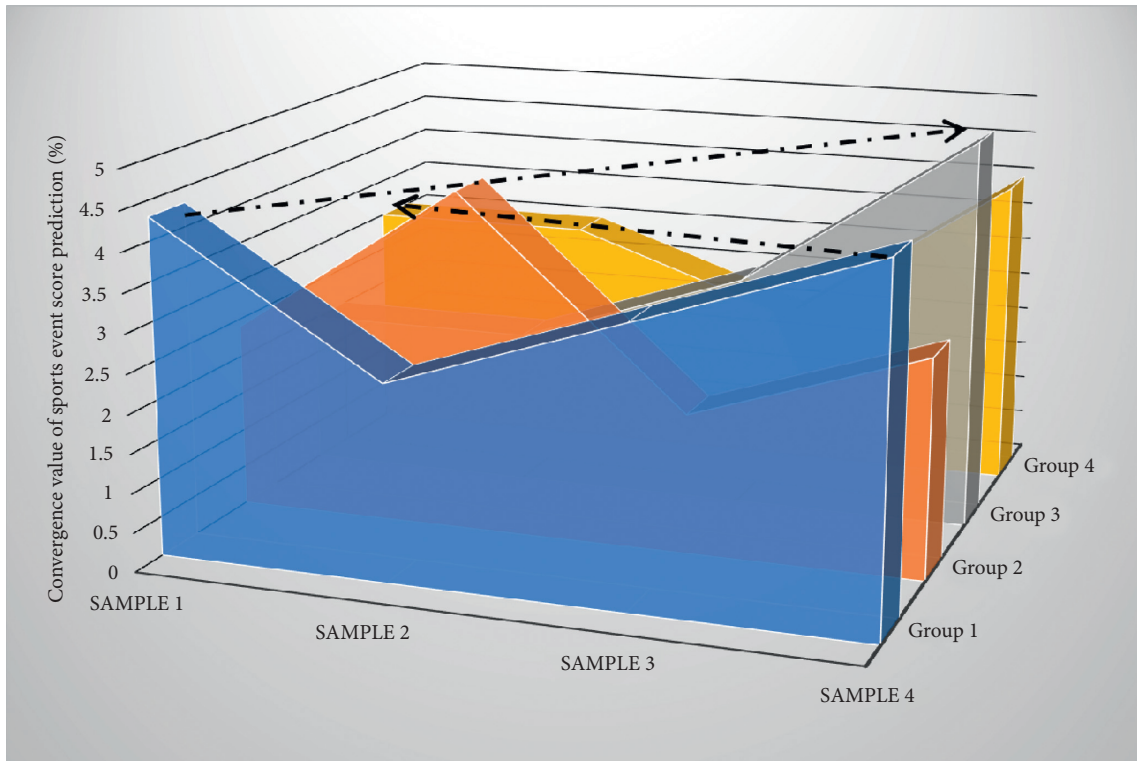
Figure 8: Convergence of the sports event score prediction network.

variables. Therefore, only the first six principal components need to be extracted under the condition that the loss of the original data is not too much information, so as to achieve the purpose of principal component extraction. It can be seen that the prediction points with larger relative errors are significantly reduced, which is 0.5% less than the maximum relative error of the nonlinear neural network, After taking the absolute value of 48 relative errors, 4 points have an error of more than 5%, 4 points are between 4% and 5%, 3 points are between 3% and 4%, and 8 points are between 2% and 2%, and the improved BP neural network has achieved better prediction results.

5. Conclusion

By reviewing relevant materials and literature, we outline data prediction and neural network related research. The artificial neural network data prediction, data transformation, data representation, and data preprocessing are studied. We made a more in-depth study on the use of artificial neural network methods for clustering, based on the in-depth study of the SOM clustering algorithm, optimized the SOM learning algorithm, and then studied the layer expansion self-organizing mapping network (GHSOM), and put forward the GRAGHSOM algorithm that introduces gray relational analysis. Regional competitiveness is a typical nonlinear complex giant system, and it is difficult for traditional linear methods to obtain satisfactory prediction results. At the same time, the experiment proves that this model has great advantages in pattern recognition. The hidden layer quantum neuron of the nonlinear model uses the linear superposition of the nonlinear basis function as

the activation function. At the same time, the thesis gives the learning algorithm of a nonlinear neural network. However, in feature selection, combining variables that are highly correlated with categorical variables may not necessarily improve the performance of the classifier. The neural network method can handle systems that are difficult to describe with mathematical models. It has strong parallel processing, self-adaptation, self-organization, and arbitrary approximation to nonlinearity. The experimental results show that the GRAGHSOM algorithm reflects the importance of each component of the sample vector in the model in the process of high-dimensional data clustering and can perform clustering more accurately; for the application of local sports score prediction (LWF) in training, we studied the perceptron neural network and successfully applied the perceptron neural network to the local sports event score prediction system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no potential conflicts of interest in this article.

Acknowledgments

References

[1] W. Zhang, R. Zhang, and R. Shang, "Application of natural computation inspired method in community detection," *Physica A: Statistical Mechanics and Its Applications*, vol. 515, pp. 130–150, 2019.

[2] L. Ruan, C. Li, Y. Zhang, and H. Wang, "Soft computing model based financial aware spatiotemporal social network analysis and visualization for smart cities," *Computers, Environment and Urban Systems*, vol. 77, Article ID 101268, 2019.

[3] P. J. Ramos-Villagrasa, P. Marques-Quinteiro, J. Navarro, and R. Rico, "Teams as complex adaptive systems: reviewing 17 years of research," *Small Group Research*, vol. 49, no. 2, pp. 135–176, 2018.

[4] Y. Hao, Y. Miao, M. Chen, H. Gharavi, and V. C. M. Leung, "6G cognitive information theory: a mailbox perspective," *Big Data and Cognitive Computing*, vol. 5, no. 4, p. 56, 2021.

[5] L. Metcalf, D. A. Askay, and L. B. Rosenberg, "Keeping humans in the loop: pooling knowledge through artificial swarm intelligence to improve business decision making," *California Management Review*, vol. 61, no. 4, pp. 84–109, 2019.

[6] R. Hristovski and N. Balagué, "Theory of cooperative-competitive intelligence: principles, research directions, and applications," *Frontiers in Psychology*, vol. 11, p. 2220, 2020.

[7] F. Kong, J. Li, and Y. Wang, "Human-computer interactive teaching model based on fuzzy set and BP neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 1, pp. 103–113, 2019.

[8] X. Xu, Z. Ye, J. Li, and M. Xu, "Understanding the usage patterns of bicycle-sharing systems to predict users' demand: a case study in Wenzhou, China," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 9892134, 21 pages, 2018.

[9] C. Schmit and J. Brisswalter, "Executive functioning during prolonged exercise: a fatigue-based neurocognitive perspective," *International Review of Sport and Exercise Psychology*, vol. 13, no. 1, pp. 21–39, 2020.

[10] O. Maksakova and L. Zhavoronkova, "Transdisciplinary team in rehabilitation of unconscious brain-damaged persons: grounds and practice," *Journal of Behavioral and Brain Science*, vol. 10, no. 8, pp. 323–343, 2020.

[11] C. Torrents, N. Balagué, Á Ric, and R. Hristovski, "The motor creativity paradox: constraining to release degrees of freedom," *Psychology of Aesthetics, Creativity, and the Arts*, vol. 15, no. 2, pp. 340–351, 2020.

[12] S. Grossberg, "Attention: multiple types, brain resonances, psychological functions, and conscious states," *Journal of Integrative Neuroscience*, vol. 20, no. 1, pp. 197–232, 2021.

[13] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2923–2960, 2018.

[14] C. Park, "Dynamic coordination phase and joint profiles: cluster and fixed-point shift techniques," *International Journal of Applied Sports Sciences*, vol. 33, no. 1, pp. 35–38, 2021.

[15] H. Yuan and G. Li, "A survey of traffic prediction: from spatio-temporal data to intelligent transportation," *Data Science and Engineering*, vol. 6, no. 1, pp. 63–85, 2021.

[16] S. Gupta, M. Kaur, S. Lakra, and Y. Dixit, "A comparative theoretical and empirical analysis of machine learning algorithms," *Webology*, vol. 17, no. 1, pp. 377–397, 2020.

[17] R. Yang and D. Wang, "Hierarchical aggregation for reputation feedback of services networks," *Mathematical Problems in Engineering*, vol. 2020, Article ID 3748383, 12 pages, 2020.

[18] C. Calvo Tapia, J. A. Villacorta-Atienza, S. Díez-Hermano et al., "Semantic knowledge representation for strategic interactions in dynamic situations," *Frontiers in Neurorobotics*, vol. 14, p. 4, 2020.

[19] T. Buszard, A. Garofolini, M. Reid, D. Farrow, L. Oppici, and D. Whiteside, "Scaling sports equipment for children promotes functional movement variability," *Scientific reports*, vol. 10, no. 1, pp. 5–8, 2020.

[20] E. Bolfíková, I. Pirohová, and M. Lenhardtová, "Bureaucracy and creative complexity–empirical analysis of problem solving effectiveness within organizations," *Sociologija i Prostor: Časopis Za Istraživanje Prostornoga I Sociokulturnog Razvoja*, vol. 57, pp. 65–91, 2019.

Research Article

Big Data Precision Marketing Approach under IoT Cloud Platform Information Mining

Wang Li (ID)

Business School, Xijing University, Xi'an, Shaanxi 710123, China

Correspondence should be addressed to Wang Li; 20030007@xijing.edu.cn

In this article, an in-depth study and analysis of the precision marketing approach are carried out by building an IoT cloud platform and then using the technology of big data information mining. The cloud platform uses the MySQL database combined with the MongoDB database to store the cloud platform data to ensure the correct storage of data as well as to improve the access speed of data. The storage method of IoT temporal data is optimized, and the way of storing data in time slots is used to improve the efficiency of reading large amounts of data. For the scalability of the IoT data storage system, a MongoDB database clustering scheme is designed to ensure the scalability of data storage and disaster recovery capability. The relevant theories of big data marketing are reviewed and analyzed; secondly, based on the relevant theories, combined with the author's work experience and relevant information, a comprehensive analysis and research on the current situation of big data marketing are conducted, focusing on its macro-, micro-, and industry environment. The service model combines the types of user needs, encapsulates the resources obtained by the alliance through data mining for service products, and publishes and delivers them in the form of data products. From the perspective of the development of the telecommunications industry, in terms of technology, the telecommunications industry has seen the development trend of mobile replacing fixed networks and triple play. The development of emerging technologies represented by the Internet of Things and cloud computing has also led to technological changes in the telecommunications industry. Operators are facing new development opportunities and challenges. It also divides the service mode into self-service and consulting service mode according to the different degrees of users' cognition and understanding of the service, as well as proposes standardized data mining service guarantee from two aspects: after-sales service and operation supervision. A customized data mining service is a kind of data mining service for users' personalized needs. And the intelligent data mining service guarantee is proposed from two aspects of multicase experience integration and group intelligence. In the empirical research part, the big data alliance in Big Data Industry Alliance, which provides data mining service as the main business, is selected as the research object, and the data mining service model of the big data alliance proposed in this article is applied to the actual alliance to verify the scientific and rationality of the data mining service model and improve the data mining service model management system.

1. Introduction

Industrial integration refers to the dynamic development process in which different industries or different sectors of the same industry penetrate and cross each other and merge into a whole, gradually developing into a new industry. Industrial integration is not only studied as a development trend, currently, but industrial integration has also developed into a realistic choice for the development of the communications industry [1]. From the perspective of the development of the telecommunications industry, in terms

of technology, there is a trend of mobile substitution for fixed networks and triple-play convergence in the telecommunications industry, and the development of emerging technologies represented by the Internet of Things and cloud computing has also led to technological changes in the telecommunications industry, making telecommunications operators face new development opportunities and challenges. In terms of business, full-service operations have been used as the main mode of operation in the telecom industry around the world [2]. Big data trading platforms led by data production or service enterprises have become the

main medium for data trading activities, and these enterprises have actively explored in specific fields and industries and formulated data trading rules, norms, and targets for their respective platforms, but there are generally unclear boundaries in functional positioning, forming multiple-segmented trading markets, making it difficult to achieve scale and industrial development, leading to low mobility of data resources [3]. This requires enterprises in different fields to integrate or cooperate as well. These companies have made active explorations in specific fields and industries, and formulated data transaction rules, specifications, and goals for their respective platforms. However, there are generally unclear boundaries in functional positioning, forming multiple-segmented trading markets, and it is difficult to achieve scale. The development of globalization and industrialization has led to low data resource mobility. Although enterprises with strong capital can supplement the data by acquiring small enterprises, one enterprise cannot achieve the comprehensive data required by market demand anyway, not to mention that acquisition activities are also unattainable for many SMEs. Therefore, many geographically dispersed data resource-based enterprises, technology-owned enterprises, service-application enterprises, network business enterprises, infrastructure enterprises, etc. The big data industry chain relies on advanced Internet and mobile Internet and other communication methods to provide data services as the mainline, by bridging the data resource gap of alliance members, in data transactions and big data industry operations, to complement each other with resource advantages [4]. By bridging the data resource gaps of the alliance members, the alliance aims at complementing resource advantages, innovating business service models, and enhancing their competitive advantages in data trading and big data industry operation, and finally forms a new cross-enterprise organization form of benefit sharing and risk sharing—big data alliance.

Although the IoT public cloud platform is developing fast and provides comprehensive functions, public cloud can support more user access and reduce cloud platform load, the public cloud access is mainly HTTP, HTTP is suitable for some devices with low sensitivity to real time, and the devices with high sensitivity to real-time are not suitable for accessing the public cloud, such as logistics monitoring in industrial production, which requires real-time monitoring logistics information flow, and when you need to control or send orders to a certain aspect of the logistics chain, you need very good real time to ensure that the orders are delivered accurately and on time. And in some enterprises, enterprises have high-security requirements for data in production, such as security monitoring, and some design enterprise originality and commercial secrets of data cannot be stored in the public cloud platform, such as the collection of enterprise equipment data, and enterprises need their servers to save equipment data [5]. What enterprises need is a platform that can meet the production needs of enterprises, especially for domestic small and microenterprises, enterprises focus on different directions, the product data in the enterprise are also

different, the demand for the platform is also different, so the public cloud platform is not suitable for diverse needs. For different data needs of different enterprises, scalability should be built to meet the needs of enterprise data platform, so a private platform for different enterprises with different data requirements should be built, a scalable platform should be built to meet the data requirements of enterprises, so building a private platform for enterprise IoT is more capable of promoting the combination of IoT technology and actual enterprise production [6].

The construction of the data aggregation service model of the original-level data aggregation service, the feature-level data aggregation service, and the decision-level data aggregation service of the big data alliance enables the members of the Alliance to break through their development bottlenecks and form a value-added industrial chain with the sharing of data resources as the cooperation link and promotes the virtualized cross-regional operation of the Big Data Alliance; on the other hand, with the data warehouse as the centre, relying on the Big Data Alliance data aggregation service model, it can provide convenient and real-time quality data services for the majority of users. Therefore, the data aggregation service model of building high-quality original-level data aggregation service, feature-level data aggregation service, and decision-making-level data aggregation service for big data alliance can rapidly accelerate the efficiency of using and sharing data resources of big data alliance and, finally, realize the value-added of data resources of big data alliance. Taking data services as the main line, by making up for the data resource gap of alliance members, in the operation of data transactions and big data industry, with the goal of complementing resource advantages, innovating business service models, and enhancing its own competitive advantages, it has finally formed benefit sharing and risk sharing a new cross-enterprise organization form.

2. Related Works

Nowadays, the competition between the major giants in the field of e-commerce is becoming more and more fierce, and in the way of marketing coordination, the effective self-operated coordination and precision marketing strategy has been more widely used and achieved certain results. As an emerging industry and tool, with the rise of e-commerce, precision marketing is also a new thing in the advertising industry, and the research results are not supported by specific systems [7]. But it is undeniable that precision marketing in today's new network business model is a more reduced cost, can bring accurate traffic means of communication, can be for small- and medium-sized enterprises in the industry in the case of tall buildings for more benefits, and have their foothold [8]. For large enterprises, accurate customer can greatly save publicity costs, and operating costs can make the enterprise to the customer's service more personalized. Future-oriented precision marketing is coming soon due to the role of capturing people's needs in the mainstay of the advertising industry [9]. With the rapid spread of information, the concept of precision marketing

has become familiar and is gradually developing into a new marketing concept. For this reason, based on the background of the rapid development of e-commerce, the connotation of personalized precision marketing and its characteristics will be targeted research, and its marketing strategy will be analyzed and discussed [10]. Personalized precision marketing has several characteristics and advantages, and the main strategies for e-commerce enterprises to carry out personalized precision marketing are to improve the level of personalized services, accelerate the process of enterprise informatization, and develop the precise positioning of enterprises [11].

The target market is the market demand that the enterprise must meet in the process of marketing, that is, the object of the enterprise's service [12]. No enterprise or product can meet the overall needs of all consumer markets, so it can only choose a certain range of target markets for the enterprise as well, and market positioning is a certain market position that the enterprise specifies to highlight its products according to the consumer demand level. After selecting the target market, due to market competition and a wide range of information, to gain market advantage, the enterprise must form a product with different characteristics from its competitors after understanding their characteristics to form a strong impression in the minds of consumers. The purpose of market positioning is to gain a competitive advantage and attract more potential consumers. This theory is very applicable to operator big data marketing [13]. This is because the marketing of operators' big data is not only to sell the operators' data to the target customers after "desensitization," but also to bundle the operators' big data with the customers' interest chain to achieve a solution where everyone benefits, and as a communication service provider, good service is crucial to the development of big data business [14].

Big data marketing generally includes four processes such as data collection, data mining, and analysis, predictive analysis, and data result in feedback. Compared with traditional marketing methods, big data marketing has advantages such as high timeliness, individual customization, high-cost effectiveness, and relevance promotion. The advent of the era of big data has brought great changes to social production and life. As a representative innovative technology, the mobile Internet has brought massive user behavior data, providing unprecedented space and potential for industries and enterprises to obtain deeper and more comprehensive insights, which makes big data marketing break through the boundaries of Internet companies, so that all walks of life can have the opportunity to have an in-depth understanding of their market environment, service targets, customer groups, or even competitors. The economic environment is an important factor in realizing demand. The operation and development of the economic environment will directly or indirectly affect the market decisions of enterprises, and it is mainly expressed in marketing as the level of economic development of a country or region and the purchasing power of consumers, which determines the size of the enterprise market.

3. IoT Cloud Platform Big Data Information Mining Precision Marketing Analysis

3.1. IoT Cloud Big Data Platform Construction. The overall architecture of the enterprise-oriented IoT cloud platform system needs to be designed from the perspective of data and enterprise users. For data, the flow between the data from the collection device to the storage system needs to be considered, and for enterprise users, the cloud platform only needs to provide users with the services they need, and the internal implementation of the cloud platform and the underlying data collection is transparent to the users [15]. Therefore, the architecture needs to meet the function of each subsystem, determine the responsibilities between each subsystem and the interface definition of other subsystems, and conform to the principle of high cohesion and low coupling defined by software engineering, and the overall architecture is easy to maintain and expand later. For example, security monitoring, and some design companies' original and commercial secret data cannot be stored on public cloud platforms. For example, corporate equipment collects data, and companies need their own servers to store equipment data. The system provides relevant services in the form of interfaces and only needs to explain the interface-related use issues and use methods, and the internal implementation logic is transparent to the user, and the user only needs to operate in the local client to complete the relevant tasks.

Data communication is the foundation of the enterprise private cloud platform and the starting point of the cloud platform architecture. In enterprise production, due to the wide variety of hardware, each hardware specification is different, communication format and device access are different, such as vehicle positioning equipment can directly connect to the open interface of the cloud platform to transmit data, but some small embedded devices such as GPS equipment cannot directly use the HTTP protocol to transmit data, so it needs embedded gateway as middleware to transmit data. Due to the different kinds of equipment measurement data and the different functions of each device, the data reporting period of some devices is very long; for example, the level meter in safety monitoring only needs to report one data a day, which does not require high real time, but, for example, in the MES workshop intelligent manufacturing system, it needs to monitor the state of the material transportation process in real time, so it needs to upload data every minute or even every 30 S or shorter period. This kind of equipment has high requirements for real-time data and for real-time commands issued by the cloud platform, as shown in Figure 1.

For the different data requirements of different enterprises, a scalable platform that meets the data needs of the enterprise should be built. Therefore, building a private platform for enterprise Internet of Things can further promote the combination of Internet of Things technology and actual enterprise production. The enterprise private IoT cloud platform needs to be able to be compatible with the connection of different types of devices with different real-time requirements. The whole communication system has
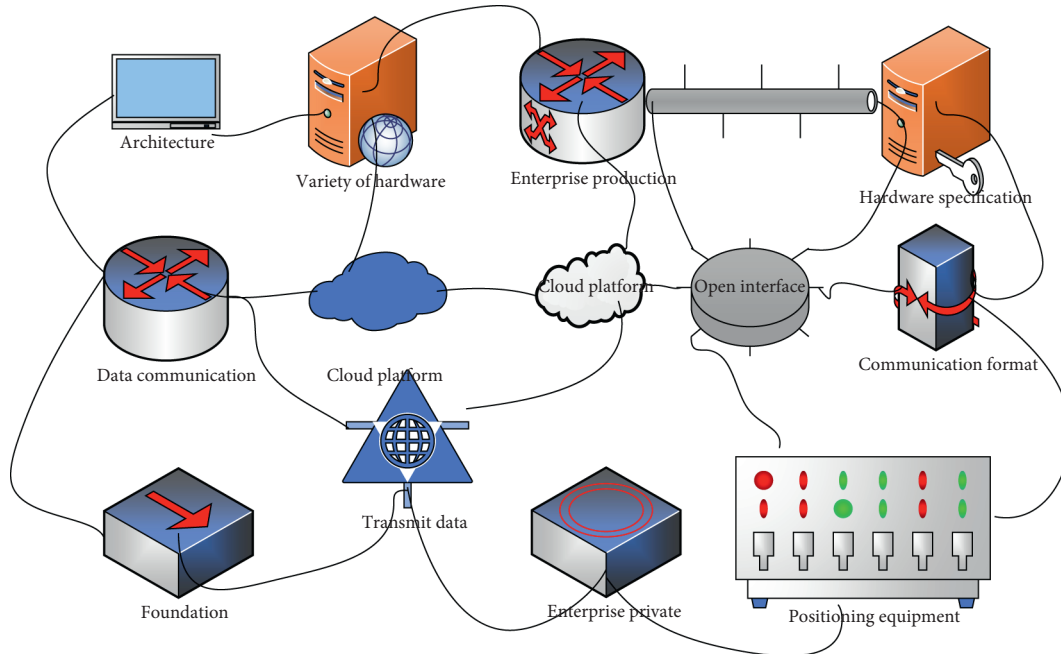
FIGURE 1: Structure of communication services.

three layers: the bottom device and embedded gateway, the intermediate device access interface layer, and the cloud platform data processing layer. For the device layer and the interface layer, due to the different real-time requirements of the devices, the cloud platform provides two types of device access: HTTP protocol access, and TCP protocol access. Since the cloud platform must support many devices, to improve the device access capability of the cloud platform, the combination of HTTP and TCP is used to complete device access. First, for devices with high real-time requirements, the cloud platform provides a real-time device access interface, and devices can directly connect to the cloud platform through the TCP to directly transmit data, and the cloud platform can actively send commands to the devices to ensure the real-time data upload and command delivery [16]. Netty is a highly available multiplexed network framework, and the underlying poll implementation can support highly concurrent connections and data exchange to improve the efficiency and speed of the system.

The data storage subsystem is built by combining the relational database MySQL with the nonrelational data available MongoDB cluster, which ensures that all data can be efficiently stored and queried through a unified data access layer to access different formats and types of data. The system uses the MySQL database for the storage of structured data, such as user data, application data, device data, system log data, and device configuration, which is used to realize the structured storage and fast access of correlated data. As the cloud platform accesses different types of devices, the data format and data dimensions collected by the devices are different, and it is difficult to store unstructured data with a relational database, the system selects MongoDB for unstructured IoT data with the simple data model, strong flexibility, huge data volume, and high requirements for data access. It can quickly accelerate the

use efficiency and sharing efficiency of data resources of the big data alliance and ultimately realize the value-added of the data resources of the big data alliance.

The cloud platform needs to provide enterprises with a range of data-based services to visualize and analyze data, using Echarte as an analysis tool to visualize and analyze data in various ways. Use Baidu Map API to provide location services for users with location needs. The platform open development interface, the caller can cognize the function of the interface role, convenient to call the development of personalized applications, as shown in Figure 2.

The IoT cloud platform consists of a communication server cluster, management server cluster, MongoDB cluster, and MySQL cluster; the communication server cluster is responsible for accessing the underlying devices, processing device uploads, processing user commands, pushing real-time data, and providing application development interfaces. The MongoDB database cluster is responsible for accessing the uploaded data of devices, and the MySQL database cluster is responsible for storing relational data of users and devices in the cloud platform.

After the device establishes a connection with the gateway, the acquisition module of the device is in the dormant state, the device does not upload data and sends heartbeat information to the gateway every time, the gateway receives the heartbeat information and parses the heartbeat, and uploads it to the cloud platform after the cloud platform receives the heartbeat information, it does not query the issued command of the device in the pending issued command table, the cloud platform replies to the gateway with the heartbeat confirmation, and the gateway receives the heartbeat confirmation. After receiving the heartbeat confirmation, the gateway sends an acknowledgment frame to the device, and the device is disconnected from the
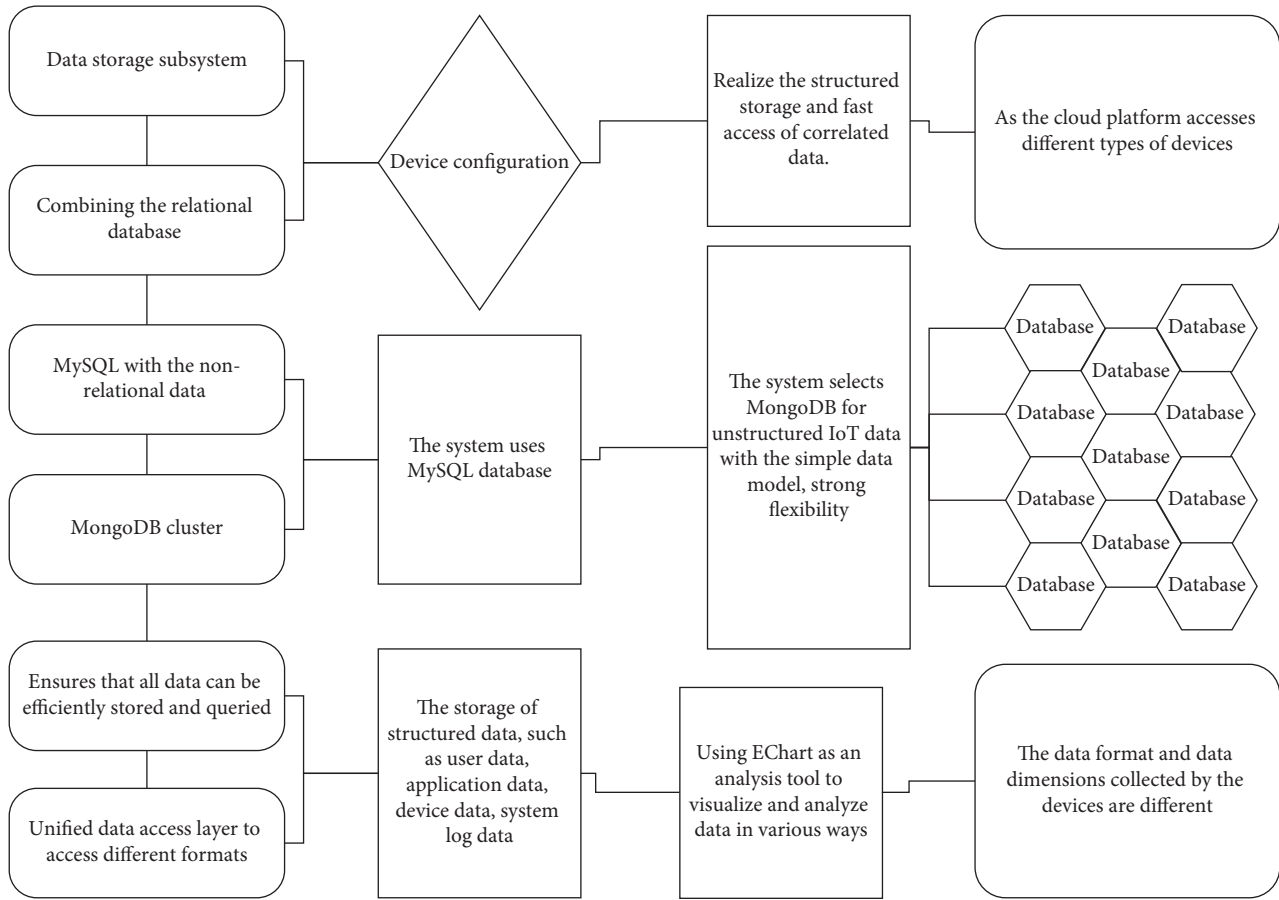
Figure 2: IoT cloud platform cluster solution diagram.

gateway. When the device acquisition module starts to collect data, the device sends the original acquisition data to the gateway, the gateway parses and re-encapsulates the acquisition data and sends it to the cloud platform, after the cloud platform receives the acquisition data and does not query the down command of the device in the pending command table, the cloud platform returns the data confirmation to the gateway, the gateway sends the data confirmation frame to the device after receiving the data confirmation, and the device is disconnected from the gateway.

When the acquisition module is collecting data, the cloud platform sends commands to the device. The device sends raw data to the gateway at regular intervals, the gateway forwards it to the cloud platform, the cloud platform receives the acquisition data and brings back the command sent by the management system to the gateway, the gateway receives the command and encapsulates the command into a command frame and sends it to the device, the device confirms the command after receiving it and sends the command confirmation frame to close the connection, the gateway sends the command confirmation message to the cloud platform, and the cloud platform transfers the command from the to-be-sent. The cloud platform transfers the commands from the command table to the already sent command table after receiving the command confirmation information.

The concept of putting the customer first has swept the world, and we are increasingly looking at what the customer thinks. Precision marketing is essentially the optimization and upgrading of segmented communication, and this model is beginning to be favoured by business owners and advertising agencies [17]. The development of precision marketing has gone through three cycles: the first stage is defined as the targeting of regional advertising messages; the second stage is the targeting of personalized customer preferences; and the third stage is the development of mobile network technology and the establishment of a personal database of Internet users, which collects and establishes a database and analyses the personalized needs of users, and then uses programmatic means to bid for advertising space in real-time, to present more precise advertising services to the public. The third stage is the development of mobile web technology and the establishment of a personalized database of web users, collecting and building the database and analyzing their personalized needs.

3.2. Information Mining Precision Marketing Design. The tag system refers to the formation of a user tag system based on basic user data through aggregation and inductive analysis, which is not directly operated externally but only used as

part of the data source for precision marketing products and other products such as the capability open platform. Data capability open platform is based on the demand of partners for data and platform resources, for partners to share Unicom data resources and infrastructure resources in a multitenant secure isolated way by opening the sample data with user identification information removed in the Unicom environment. Conduct in-depth discussions and research, and formulate corresponding marketing improvement countermeasures, which have a certain enlightening effect on the company's optimization of marketing strategies. Digital marketing is relying on massive data and big data processing capability, and under the premise of ensuring data privacy and security, optimizing the placement strategy and channels for customers in various industries through in-depth mining and analysis of big data, providing accurate marketing reach means to achieve the purpose of reducing marketing costs, and improving marketing effectiveness. Smart Footprint is to provide big data analysis services of user groups such as population density, cross-domain migration, and travel analysis for urban planning, traffic planning, and outlet location by pinpointing the number of populations within a certain granularity.

Internet data are limited by its own data genes, which are manifested in the following four aspects: first, the data of Internet companies are fragmented, e-commerce only has e-commerce sales data, no Baidu search data, and the data exists locally; second, due to company strategy and data security requirements, few Internet companies are willing to open their data to partners or competitors. Even if there is openness, more are only at the business model level and application level, and the data level is still relatively closed; third, Internet data integration is difficult, and the registered personal account is also short-term, and the account information is unstable; and fourth, the business field of Internet companies directly determines the content of data, and the scope of data is not comprehensive. Data flow requires multipart cooperation, so the architecture needs to clarify the functions between each subsystem, determine the responsibilities between each subsystem, and define the interface for other subsystems.

$$u_A = qk_0 (\beta_A - \delta_A \xi_{BA} - \delta\beta\alpha_{AB} + q_0 k_0 \zeta_{AB}),$$
$$u_A' = qk_0 (\beta_A + \delta_A \xi_{BA} - \delta\beta\alpha_{AB} - q_0 k_0 \zeta_{AB} + w). \tag{1}$$

Through the division of labour and cooperation, member companies in the big data alliance jointly develop the technologies required in the field of big data processing and provide data service solutions for users, and data resource sharing runs through the entire data service-application development process of the alliance [18]. Combined with the previous analysis process of the equilibrium state of data resource sharing in a big data alliance, the growth process of data resource stock of each member under the influence of the data resource sharing coefficient is shown in Figure 3. The results show that the data resource sharing coefficient shows a positive feedback relationship with the equilibrium state of data resource
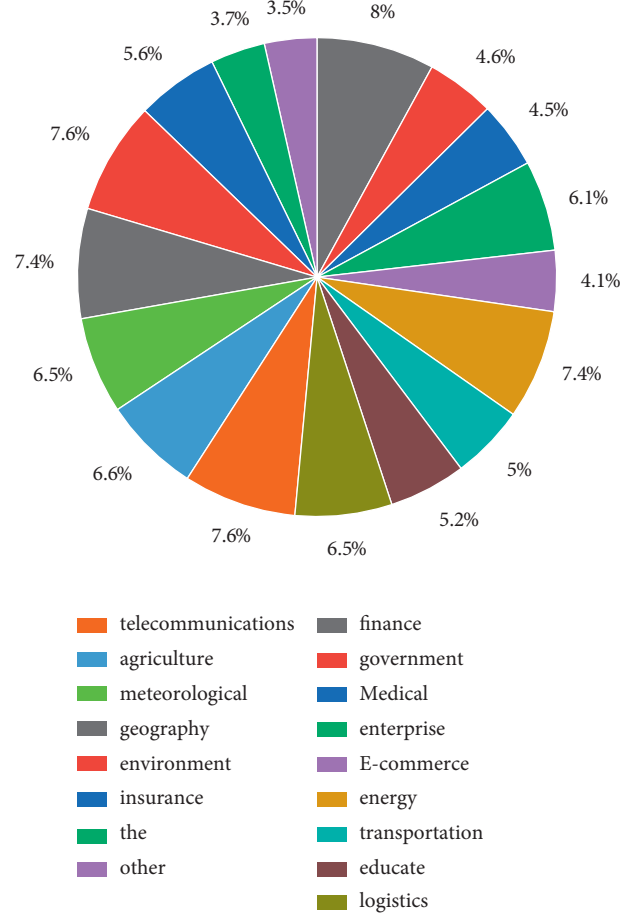
FIGURE 3: Market share of big data applications by industry.

sharing, which directly affects the equilibrium point values under different sharing relationship patterns. By adjusting and coordinating the coalition data resource sharing coefficients, the data resource stock growth will show different equilibrium states.

Data aggregation refers to the use of scientific aggregation methods to orderly connect the scattered and isolated data resources scattered among different members of the big data alliance in a certain way, and implement the structural system rebuilding, so that data resources from different sources, at different levels, and with different contents can be integrated comprehensively, thus forming a valuable and complete system. Data correlation refers to the value transformation process of data to information (knowledge) by extracting information or extracting some knowledge from data resources by organizing and extracting features in data resources based on correlation analysis, and achieving some interesting discoveries. Data fusion refers to the process of conversion, integration, and merging in information or knowledge by using advanced big data processing technology to further concentrate and refine to form new information and new knowledge, to construct effective information resources and knowledge resources for users' domain problems-solving and release the deep potential value of big data:

$$\text{POS}_p(Q)_{(x)} = \cup PX(x^2),$$

$$\prod_B^A A(I_1, I_2, \ldots, I_m) = \bigcap_{a_j \in B} Ajl_j^2. \tag{2}$$

From the perspective of value cocreation, data aggregation is the data aggregation at the original level of the big data alliance, the data value breeding stage. In the very early days, data were considered valuable, but data that could breed great business value were not data in the traditional sense but should be referred to as data elements or data resources. This is because while data contain a great value, it needs to be further developed and utilized before the value of the data can be gradually released and amplified, which is essentially like elements and resources. In the operation of a big data alliance, whether the data resources scattered in different channels correctly and effectively brought together in the alliance is the key to the aggregation of data at the original level, which will directly affect the efficiency of the alliance operation and the final service results. It is necessary to consider the overall performance loss caused by the mutual calls between subsystems and the flow of data between the various subsystems, as well as the complexity of the cloud platform in the development process:

$$M_B = \{D_k^2 | D_B^2\},$$

$$y(x) = \text{sgn} \sum_{k=1}^N a_k \varphi(x, x_k^2). \tag{3}$$

In general, core enterprises in big data alliances tend to be more willing to proactively share data resources to maintain their core position in the network, and noncore enterprises mainly play the role of supporting support. Data sharing will be more likely to occur when a mutually beneficial symbiotic relationship is formed among the member enterprises of the alliance. In particular, the more the values of different types of member enterprises in the alliance converge, the easier it will be to understand and trust each other in cooperation, which is conducive to forming a consensus and sense of identity for data sharing:

$$F = \arg \min(M_{K \times H} g_H^2). \tag{4}$$

A service tagging system is a tag library built around multiple entity objects and connections between entities using the method of tagging description, which is an important foundation for building portraits, graphs, and networks [19]. The big data alliance establishes a rich tag library through a combination of manual and machine tagging methods while monitoring the use and effects of various types of tags, counting popular and selected tags in real time, replacing unreasonable tags in the library, and continuously adding new tags according to the needs of the business to further promote the improvement of the big data alliance data tagging system and form a dynamic collection of data interpenetration-related service tags. At the same time, if users have many instances of data resources, they can support the classification management of instance resources on the platform by binding tags to the instance resources, so

that users can customize the new tags, as shown in Figure 4. It includes the underlying device and embedded gateway, the access interface layer of the middle device, and the data processing layer of the cloud platform.

In the process of feature-level data aggregation services, the complex data structure at the bottom of the alliance is separated by the tag model view layer to achieve the flexible configuration of data resources, which can greatly enhance the operational experience of member enterprises in specific situations. Specifically, we try to flexibly build the required business modules on top of the tags, use the tags as tools to realize unified business computing operations on data, and support on-demand tag selection according to the needs of business scenarios. In the actual operation process, under the premise that the big data alliance needs to integrate the data resources of multiple members, this dynamic tag extraction approach has very good scalability [20]. In the whole service panorama modelling process, due to the limitations and inadequacies of a single method, the cross-fertilization of a rich system of multidisciplinary fusion methods is needed to support the flexible extraction of features from these massive data resources and to realize the reasonable allocation and scheduling of service resources in the service panorama modelling process to quickly respond to the personalized needs of users. big data alliance provides the function of sharing the same label attempt between multiple modules so that the results of data interpenetration correlation analysis can be further combined with algorithm module, tool module, and method module, and the objects screened by the analysis can be combined with other application systems, such as combining multiple application scenarios such as operation and maintenance management of intelligent transportation system, customer relationship management in the marketing field, and equipment supervision of enterprise production and operation. Help users to solve the core problems in specific field contexts, and improve their data service satisfaction.

4. Results and Analysis

4.1. IoT Cloud Big Data Platform Performance. The shard mode is a way to slice the data horizontally, it includes a routing server mango, a configuration server mongo, and a slice server shard the data are stored by the slice server, the routing server is the entrance of the whole database cluster, and all read and write requests have to go through the routing server to distribute the requests to the corresponding. All read and write requests are distributed to the corresponding shard server. By directly transmitting data, the cloud platform can actively send commands to the device, ensuring the real-time data upload and command issuance. The TCP access method is built using the Netty framework. Netty is a highly available multiplexed network framework. The bottom layer is implemented using epoll, which can support high concurrent connections and data exchanges, and improve the efficiency and speed of the system. The configuration server stores the routing slice information of the database, and the routing server itself does not store information but reads the cluster
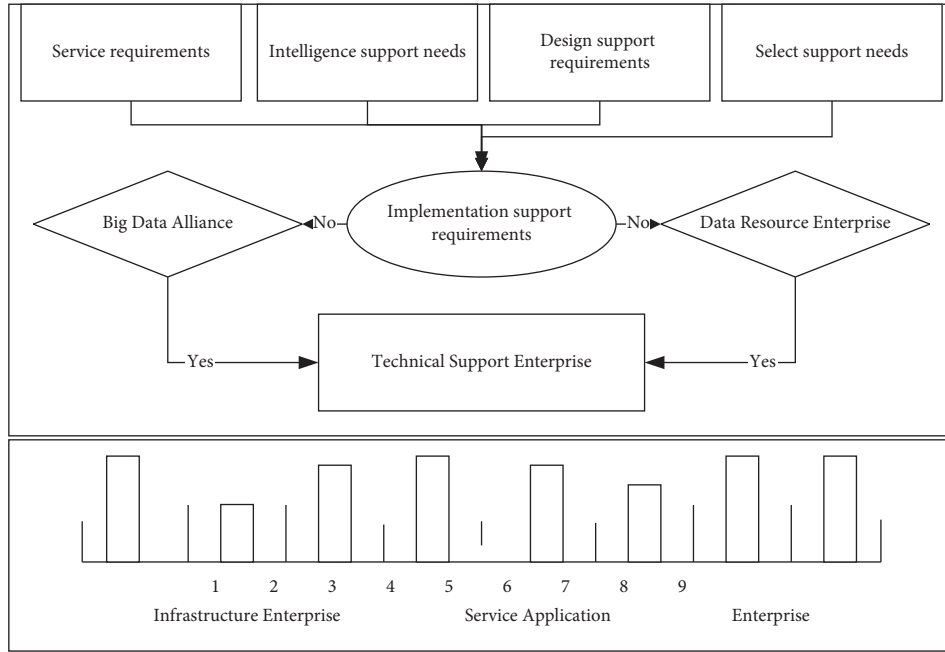
Figure 4: Schematic diagram of the formation of a decision-level data aggregation service offering.

configuration from the configuration server into memory at start-up. The slice server is responsible for processing the requests forwarded by the routing server to avoid excessive CPU, memory, and I/O pressure on one server. To ensure data storage security, efficient read and write, and storage system scalability, the system uses a combination of slicing and replica sets, which can ensure the data disaster recovery capability, but also enable the separation of reading and write, while the slicing mode has high scalability, can be extended to cloud computing and other data processing methods, as shown in Figure 5.

As can be seen, by the real-time analysis chart of the cloud platform system, the system can analyze the real-time fluctuation of the data to the size of the data. The data interface also analyses the data in terms of location, and the test device uploads the device coordinates, which need to be converted into the standard coordinates used by the map to improve the accuracy of the display when using Baidu Maps to display the location. The performance of the IoT cloud platform is related to the quality of the services provided by the cloud platform, which needs to be able to run stably under the pressure of high concurrency. The performance of the cloud platform is tested using the JMeter tool, which is an efficient stress testing tool based on Java technology and can simulate huge concurrency to stress test the server. In this article, we first conducted stress tests on the HTTP-based device access interface. In the basic configuration, the connection protocol is HTTP, the system test device access interface listens to port 80, the interface IP address is 192.168.3.232, the interface path is data, the device access interface sends data using POST request to send data, the data are in JSON format, the stress test data include the id of the simulated device, and the simulated data the data include temperature, humidity, latitude, and longitude.
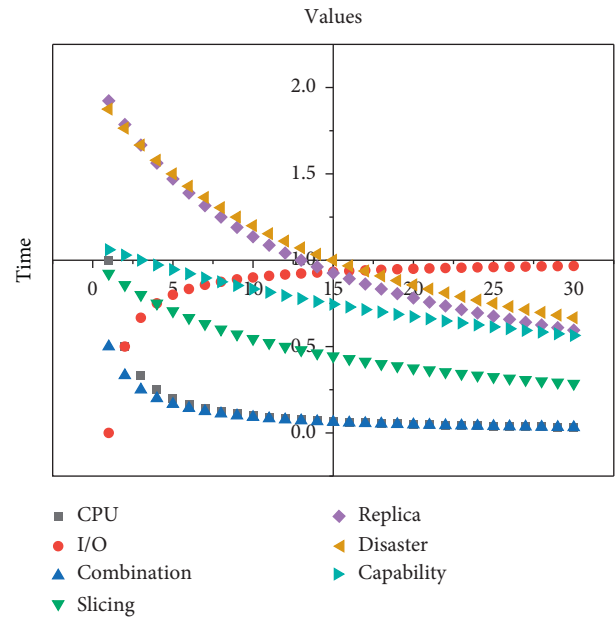
Figure 5: Real-time analyses of data from the cloud platform.

From the test results of the cloud platform, we can see that the maximum response time of the system is 266 ms, the minimum response time is 21 ms, and the average response time is 50 ms. For actual use, due to the equipment conditions and network conditions, the equipment upload cycle is greater than the maximum response time. During the stress test, the system processing error rate is 0.00%, and the request throughput of the system reaches 76.2 times per second, which means that under the operating environment of processing 76.2 requests per second, the average data processing per second reaches 219 KB, and the system is fully

capable of correctly processing the data uploaded by the device, as shown in Table 1.

It can be seen that within 100 seconds, the host accepts 4800 TCP requests and the connection is not open; it can be seen that the average processing time of the cloud platform TCP device access server for requests is 54 ms, the minimum processing time is 19 ms, and the longest processing time is 260 ms; for the high real-time requirements of the device, it can fully meet the requirements of the device for real time, and at the same time, the server throughput can reach 47.8 times per second, which means that the server simultaneously processes 47.8 requests per second and 219 bytes of data per second, the error rate of the server is 0.00%, and the system can support normal communication of the usual system in the enterprise.

The platform application module functions, device module main functions, device configuration, issuing commands, and data display module were tested to meet the functional requirements of the cloud platform analysis, and it can fulfill the requirements of equipment access, data processing, data storage, data analysis, and so on. In the performance test, the cloud platform used the JMeter stress test tool to test the data processing capability of the cloud platform in the face of large-scale device access, and the system first used JMeter to test the performance of the HTTP protocol-based device access interface; the average processing time is 50 ms, and the maximum throughput can reach 76.2/sec. Then, the real-time device access interface was tested for performance. The test results show that the processing speed and results of the cloud platform can meet the actual production needs when facing large-scale device access.

4.2. Information Mining Precision Marketing Results. In the big data industry chain, the core competencies vary due to the different positions of the alliance members and have the following characteristics: members located in the downstream of the industry chain usually have strong capabilities in data collection, processing, and handling; members located in the midstream of the industry chain usually have strong capabilities in analysis, technology, and model building; and members located in the upstream of the industry chain usually have strong comprehensive capabilities and decision analysis, and knowledge expression capabilities. Therefore, the big data alliance establishes a capacity collaborative scheduling mechanism to give full play to the capacity advantages of each member and promote the alliance can customize data mining services for users with the optimal capacity combination. The cooperative scheduling guarantee among members is an organizational form of scheduling guarantee based on resource collaborative scheduling and capability collaborative scheduling. Through collaborative scheduling among members, it facilitates experience sharing, knowledge exchange, and member interaction to stimulate group wisdom. Therefore, the big data alliance establishes the members' collaborative guarantee mechanism to realize the integration of multiple subjects' wisdom and thus create high value in the process of realizing customized data mining services.

TABLE 1: Platform performance.

Label	Sample	Average	Max
HTTP	4.41	3.18	4.75
Request	2.46	6.12	8.82
Total	5.71	8.62	8.47
Mean	3.24	8.94	8.41

When a user makes a data mining service request, their service request needs to be analyzed and service searched. Retrieve whether there are similar historical service cases in the service case library of the big data alliance data mining service platform. If similar historical service cases exist, then service acquisition needs to be performed in the selected intelligent data mining service model, while if there are no service cases like the target case in the big data alliance historical service case library, then service acquisition needs to be performed using customized data mining services, as shown in Figure 6.

Big data alliance customized data mining services are provided by data resource-based enterprises, data technology-based enterprises, and data application-based enterprises to provide users with more targeted and personalized customized data mining services when the standardized data mining service content cannot meet user needs. The gateway receives the command and encapsulates the command into a command frame and sends it to the device. The device confirms after receiving the command and sends a command confirmation frame to close the connection. The gateway sends the command confirmation message to the cloud platform. The sending order list is transferred to the already issued order list.

Firstly, the guiding role of user-user needs should be clarified, and the construction of a feedback system for data aggregation services at the decision-making level should be emphasized as the key to the orderly development of the big data alliance. Secondly, the layout of service facilities should be scientifically formulated and reasonably planned, the allocation of service resources should be optimized, efforts should be made to cover a full range of user needs, and multilevel service feedback channels and channels should be developed. Finally, for the same user's service feedback information should pay attention to the collection from multiple channels, to form a closed-loop complete circuit of data aggregation services at the decision-making level, as shown in Figure 7.

As an emerging decision support system, the decision-level data aggregation service model extends the scope of its users to governments, enterprises, organizations, individuals, etc. Especially for enterprises, it not only supports strategic decisions of middle and senior management but also serves front-line business personnel, even including suppliers and partners. The decision-level data aggregation service model assigns service tasks to specific member enterprises through the data aggregation service platform, which requires continuous cooperation and a high degree of collaboration between different types of member enterprises so that the relevant service business content is agreed upon and the consensus is reached among the service subjects without easy ambiguity, and the joint provision of
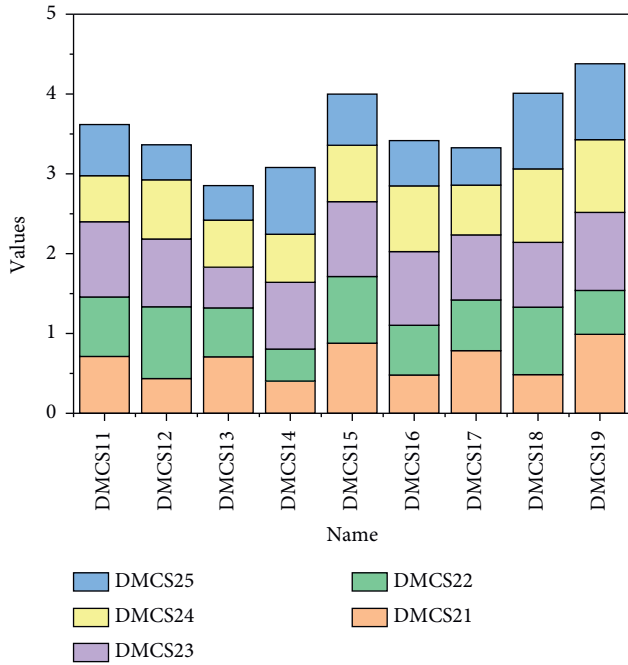
FIGURE 6: Capability vector distance.
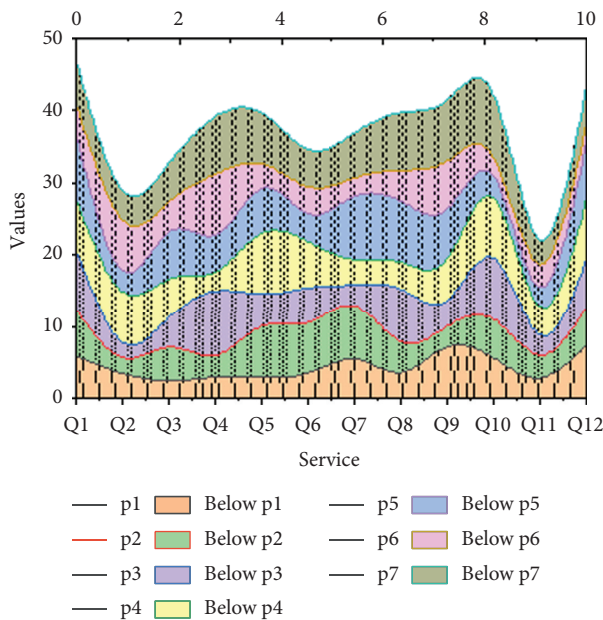
FIGURE 7: Results of the business model selection problem attributes.

intelligence support products, design support products, selection support products, and implementation support products service solutions.

5. Conclusion

For the heterogeneous characteristics of data, the combination of the MongoDB database and MySQL database is used to store data in data storage, which achieves efficient storage and fast access of massive IoT data, and optimizes the

data storage method for the temporal nature of IoT data, and stores the data in aggregation to improve the data reading speed. For the storage of massive IoT data and data security, the cloud platform designs a MongoDB database cluster for IoT data storage, which uses a combination of sharing and replica sets to improve the scalability of data storage and the processing capability of a large amount of data storage access. Through the analysis of the big data marketing environment, its advantages in big data marketing are greater than its disadvantages, and with the guidance of government policies and enterprises for their own development needs, big data marketing in Gansu also has a better external environment, to seize market opportunities through differentiated marketing strategies. Based on the problems and shortcomings of the big data marketing strategy, the target market of big data marketing is repositioned, and it is concluded that the government-industry, tourism industry, financial industry, and SME customer market are the key target markets of big data, and for these target markets, the marketing strategy that is more suitable for the long-term development of big data marketing is designed, that is, the strategies of product creation differentiation and operational process differentiation. In this way, we can enhance the comprehensive competitiveness of big data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

Acknowledgments

References

[1] B. Zhang and B. Zhang, "Precise marketing of precision marketing value chain process on the H group line based on big data," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 3, pp. 2837–2845, 2018.

[2] S. Zhao and J. Ma, "Research on precision marketing data source system based on big data," *International Journal of Advanced Media and Communication*, vol. 7, no. 2, pp. 93–100, 2017.

[3] F. Cui, H. Hu, and Y. Xie, "An intelligent optimization method of E-commerce product marketing," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4097–4110, 2021.

[4] Z. Salih Ageed, S. Zeebaree, M. Mohammed Sadeeq et al., "A survey of data mining implementation in smart city applications," *Qubahan Academic Journal*, vol. 1, no. 2, pp. 91–99, 2021.

[5] M. Yang, M. Mahmood, X. Zhou, S. Shafaq, and L. Zahid, "Design and implementation of cloud platform for intelligent logistics in the trend of intellectualization," *China Communications*, vol. 14, no. 10, pp. 180–191, 2017.

[6] T.-M. Choi, S. W. Wallace, and Y. Wang, "Big data analytics in operations management," *Production and Operations Management*, vol. 27, no. 10, pp. 1868–1883, 2018.

[7] C. Jinbo, Z. Yu, and A. Lam, "Research on monitoring platform of agricultural product circulation efficiency supported by cloud computing," *Wireless Personal Communications*, vol. 102, no. 4, pp. 3573–3587, 2018.

[8] J. Wang, Y. Yang, T. Wang, R. S. Sherratt, and J. Zhang, "Big data service architecture: a survey," *Journal of Internet Technology*, vol. 21, no. 2, pp. 393–405, 2020.

[9] V. Palanisamy and R. Thirunavukarasu, "Implications of big data analytics in developing healthcare frameworks-a review," *Journal of King Saud University - Computer and Information Sciences*, vol. 31, no. 4, pp. 415–425, 2019.

[10] G. Morota, R. V. Ventura, F. F. Silva, M. Koyama, and S. C. Fernando, "Big data analytics and precision animal agriculture symposium: machine learning and data mining advance predictive big data analysis in precision animal agriculture 1," *Journal of Animal Science*, vol. 96, no. 4, pp. 1540–1550, 2018.

[11] C. Jinbo, C. Xiangliang, F. Han-Chi, and A. Lam, "Agricultural product monitoring system supported by cloud computing," *Cluster Computing*, vol. 22, no. 4, pp. 8929–8938, 2019.

[12] C. Iorio, G. Pandolfo, A. D'Ambrosio, and R. Siciliano, "Mining big data in tourism," *Quality & Quantity*, vol. 54, no. 5, pp. 1655–1669, 2020.

[13] T. Qian, S. Zhu, and Y. Hoshida, "Use of big data in drug development for precision medicine: an update," *Expert review of precision medicine and drug development*, vol. 4, no. 3, pp. 189–200, 2019.

[14] T. Guo and Y. Wang, "Big data application issues in the agricultural modernization of China," *Ekoloji*, vol. 28, no. 107, pp. 3677–3688, 2019.

[15] N. Tantalaki, S. Souravlas, and M. Roumeliotis, "Data-driven decision making in precision agriculture: the rise of big data in agricultural systems," *Journal of Agricultural & Food Information*, vol. 20, no. 4, pp. 344–380, 2019.

[16] Y. Liu, "Incomplete big data imputation mining algorithm based on BP neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 4, pp. 4457–4466, 2019.

[17] R. K. Jena, "Sentiment mining in a collaborative learning environment: capitalizing on big data," *Behaviour & Information Technology*, vol. 38, no. 9, pp. 986–1001, 2019.

[18] C. Liu, Y. Feng, D. Lin, L. Wu, and M. Guo, "Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques," *International Journal of Production Research*, vol. 58, no. 17, pp. 5113–5131, 2020.

[19] J. Y. Lee, J. S. Yoon, and B.-H. Kim, "A big data analytics platform for smart factories in small and medium-sized manufacturing enterprises: an empirical case study of a die casting factory," *International Journal of Precision Engineering and Manufacturing*, vol. 18, no. 10, pp. 1353–1361, 2017.

[20] H. Jung and K. Chung, "Social mining-based clustering process for big-data integration," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 589–600, 2021.

Research Article

E-Commerce Credit Risk Assessment Based on Fuzzy Neural Network

Lina Wang [1] and Hui Song [2]

¹School of Finacial Technology, Hebei Finance University, Baoding, Hebei 071051, China
²School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Haidian, Beijing 100876, China

Correspondence should be addressed to Lina Wang; wanglina@hbfu.edu.cn

In this paper, we propose a cooperative strategy-based self-organization mechanism to reconstruct the network. The mechanism includes a comprehensive evaluation algorithm and structure adjustment mechanism. The self-organization mechanism can be carried out simultaneously with the parameter optimization process. By calculating the similarity and independent contribution of normative neurons, the effectiveness of fuzzy rules can be jointly evaluated, and effective structural changes can be realized. Moreover, this mechanism should not set the threshold in advance in practical application. In order to optimize the parameters of SC-IR2FNN, we developed a parameter optimization mechanism based on an interaction strategy. The parameter optimization mechanism based on a joint strategy, namely multilayer optimization engine, can split SC-IR2FNN parameters into nonlinear and linear parameters for joint optimization. The nonlinear parameters are optimized by an advanced two-level algorithm, and the linear parameters are updated with the minimum biological multiplication. Two parameter optimization algorithms optimize nonlinear and linear parameters, reduce the computational complexity of SC-IR2FNN, and improve the learning rate. Using the principal component factor analysis method, seven representative common factors are selected to replace the original variables, which include the profitability factor of the financing enterprise, the solvency factor of the financing enterprise, the profitability factor of the core enterprise, the operation guarantee factor, and the growth ability of the financing enterprise. Factors, supply chain online degree factors, financing enterprise quality, and cooperation factors, can well measure the credit risk of online supply chains. The logistic model shows that the profitability factor of the financing company, the debt repayment factor of the financing company, and the profitability of the core company are three factors that have a significant impact on the credit risk of online supply chain finance. Based on the improved credit calculation model, we developed an online clue risk calculation. This method is based on site conditions and can evaluate credit risk. From the test results, the improved credit scoring system is the result of facing speculative and circular credit fraud and implies that the traders of risk commentators are in a leading position in each electronic device. The results show that risk analysis is effective in any case.

1. Introduction

At present, artificial intelligence is developing rapidly, and artificial neural network algorithm is the core of research [1]. The combination of artificial neural networks and traditional industries is an effective way to solve traditional agricultural problems. The research purpose of this subject is to combine the fuzzy logic system and the artificial neural network into the fuzzy neural network through research and use it in

e-commerce credit risk assessment [2]. Fuzzy control is to imitate people's judgment and plot reasoning to deal with problems that are difficult to solve by general methods, especially nonlinear problems, problems that cannot be modeled but require precision. The membership function and rule design in fuzzy logic control are all artificially set, and in software design, the more the rules there are, the worse the real-time performance of the control operation will be. The artificial neural network has good self-adaptive

learning ability. It mainly works by simulating the biological neural network, but the disadvantage is that its ability to express rules is poor. Therefore, the fuzzy logic system and neural network are combined to make full use of the advantages of both, make up for the shortcomings of both, and form a fuzzy neural network algorithm [3, 4].

With the development of the e-commerce industry, its risk issues have become increasingly prominent [5]. Credit risk has become an important reason hindering the development of the industry. Various disputes and complaints caused by credit risk issues reduce consumer trust and repeat purchase rates, increase customer acquisition costs and user transaction costs of e-commerce platforms, and hinder the further development of e-commerce [6]. According to the investigation of the Internet fraud observation website, in areas with mature e-commerce markets such as the United States, the amount of money lost by traders due to seller credit risk is also increasing year by year; while China's market system is imperfect, legal and credit systems are imperfect. The traders are more likely to suffer losses because of this. Relevant scholars believe that the issue of credit risk has become an industry and social problem that needs to be resolved urgently [7]. The influencing factors of e-commerce credit risk issues include technical reasons not only caused by the separation of time and space on the Internet but also related to whether the management of the virtual market is perfect and the soundness of the legal system, and the most important thing is the credit choice of the transaction subject [8]. The information asymmetry caused by the virtual nature of the network has aggravated the inequality of information between the parties to the transaction. From the perspective of the transaction characteristics of e-commerce, the e-commerce platform has large user traffic and low stickiness, which leads to short-term interests driven by sellers and triggers their actions [9, 10].

The interval class II fuzzy neural network structure asks the problems of solution difficulty and computational complexity to design the interval class II structure. The interval second kind of fuzzy Shinto network is constructed in an organized structure, and the computational complexity is low. B2C electronics led event supply finance credit risk crunch creates a comment price model and defines a comment price index by analyzing the models attributed to factor analysis and logistic back. When calculating the risk of a transaction, according to historical transaction records, the risk calculation is divided into three categories. When calculating the risk, the commodity price, seller credit value, seller credit rating, historical average price, and transaction failure threshold are considered. The calculation of the risk of cycle deception also takes into account the important factor of the deception cycle. At the same time, this article also gives the algorithm idea of seeking the deception cycle. Through experiments on periodic fraud and credit speculation transactions, we can see the impact of commodity prices, credit values, credit ratings, and transaction failure rates on the value of risk. The graphic image drawn by MATLAB illustrates the historical transaction process. The credit scoring method proposed in this article closely links to credit and risk and divides commodity prices into ranges.

The scoring of "good reviews" and "bad reviews" changes dynamically with the progress of transactions. This method effectively suppresses periodic fraud by increasing the opportunity cost of users with high credit values. Using the range occupancy rate of commodity prices to score effectively suppresses credit speculation and protects the interests of sellers who only sell low-priced commodities. The experimental results show that the scoring method is effective and feasible.

2. Related Work

Relevant scholars improved the genetic algorithm and then used the improved algorithm to optimize the weights of the feedforward network so that the training accuracy of the algorithm was significantly improved [11]. Researchers combined the traditional BP algorithm and the differential evolution algorithm to propose a new weight training method and used it in breast cancer prediction experiments and achieved good results [12]. Related scholars combined the differential evolution algorithm with EANN to construct the MPANN network model, which improved the function approximation ability [13].

Relevant scholars introduced mutation operators on the basis of the PSO algorithm and then intelligently fused with wavelet neural network, diagnosed transformer faults, and obtained good experimental results [14]. Relevant scholars optimized the convolutional neural network and proposed a convolutional neural network acceleration algorithm based on parameter and feature redundancy compression, which was used for image processing and proved the optimization performance of the algorithm [15]. Neural network technology has gradually developed and been used in fault detection, prediction, and classification.

It realizes the processing and expression of the ambiguity in data through membership functions and fuzzy neurons. Relevant scholars combine evolutionary computing with neural networks to propose multiple groups of classification algorithms to effectively adopt the membership functions of the fuzzer and defuzzer to the data set and use actual economic data to successfully test [16]. Related scholars have proposed a new online sequential learning evolution RL neuro-fuzzy model design and developed a dynamic evolutionary fuzzy neural network (DENFIS) function approaching the RL system method [17]. Relevant scholars apply a fuzzy neural network to environmental safety assessment, using fuzzy neural networks to deal with the characteristics of fuzzy phenomena, and have achieved good results [18, 19].

Relevant scholars use weighted Euclidean distance to improve traditional clustering, use rough sets to reduce attributes, build a neural network based on the clustering results, and establish a new prediction model and have achieved good results in wind speed prediction [20]. The researchers gave a theoretical overview of the fuzzy nervous system, discussed related knowledge, introduced two network models, and proved the performance of the model through simulation experiments [21]. Relevant scholars proposed that the data can be input in batches during the reasoning process, and the data generated by the overall rules

can be divided, thereby reducing the number of rules and completing the algorithm optimization of the fuzzy neural network, which makes it more advantageous in processing high-dimensional data [22]. Relevant scholars analyzed biological principles and proposed the method of using gene overlap to optimize the fuzzy neural network, generating T-S fuzzy rules through genetic code shifting, combining genetic algorithm with fuzzy control, realizing genetic mutation, and improving the performance of the algorithm [23].

Relevant scholars have incorporated average purchase prices, transaction density, and historical transaction records into the evaluation index system to comprehensively reflect the credit status of sellers in the transaction process [24]. Researchers and others introduce commodity prices and the credit of the evaluation subject into the credit evaluation model to identify whether there are false evaluations in the transaction process between buyers and sellers, thereby reducing the uncertainty of credit evaluation results [25]. Relevant scholars have proposed that an effective way to distinguish honest sellers and prevent dishonest sellers from trading is to establish an effective credit evaluation mechanism [26].

On the basis of previous studies, relevant scholars have investigated the historical transaction volume of sellers and the credit level of evaluation subjects and objects into the credit evaluation model [27]. Relevant scholars pointed out that the current credit evaluation model indicators cannot provide sufficient differentiation [28]. Therefore, indicators such as the historical credit of buyers and sellers and the number and amount of transactions should be increased to more accurately reflect the user's credit.

Researchers propose an online credit evaluation method that measures the similarity between new transactions and past transactions in the dimensions of product type, number of sold products, and transaction amount, thereby establishing a multidimensional credit evaluation index system [29]. From the perspective of evaluation semantics, relevant scholars have proposed an online credit evaluation system based on context, which aims to assist consumers in measuring the credibility of sellers and to screen whether consumers' evaluation opinions are of reference [30].

The topology of the neural network has a greater impact on its performance and calculation speed. Compared with the general neural network, the IT2FNN structure is more complex and contains more parameters and faces a greater computational burden in practical applications. In general, more neurons can ensure that the neural network has better performance, but too many neurons will make the calculation of the neural network too complicated, which is not conducive to practical applications. In addition, fewer neurons will reduce the performance of the neural network. Therefore, how to determine the appropriate IT2FNN structure has always been the focus of research.

3. Method

3.1. Interval Type 2 Fuzzy Neural Network. IT2FNN not only has excellent uncertainty processing capabilities but also has adaptive learning capabilities, so it can well realize the identification and control of complex nonlinear systems with uncertainties and time-varying properties. Different from a type 1 fuzzy neural network, IT2FNN adopts an interval type 2 membership function to convert the exact value into an interval fuzzy set so that it can better deal with the uncertainty in the system. In addition, IT2FNN avoids the use of primary and secondary membership functions in the type 2 fuzzy neural network to cause an overly complicated calculation process so that IT2FNN can be better applied in practical engineering.

The structure of IT2FNN is shown in Figure 1. Its structure specifically includes five layers of neurons, which are the input layer, subordinate layer, rule layer, subsequent layer, and output layer.

The output of IT2FNN is as follows:

$$
y(t) = y'(t)q(t+1) - y''(t)[1 - q(t+1)],
$$

$$
y'(t) = \frac{\prod_{j=0}^{M-1} f'_j(t)h_j(t+1)}{\prod_{j=0}^{M-1} f'_j(t)},
$$

$$
y''(t) = \frac{\prod_{j=0}^{M-1} f'_j(t)h_j(t+1)}{\prod_{j=0}^{M-1} f''_j(t+1)},
$$

$$
h_j(t) = \prod_{i=0}^{n-1} \left[w_{ij}(t+1)x_{i+1}(t) \right] + b_j(t).
$$

where $y'(t)$ and $y''(t)$ are the lower and upper output bounds of the consequent layer at time t, $q(t)$ is the scale factor, $h_j(t)$ is the consequence of the j-th fuzzy rule, $w_{ij}(t)$ is the consequent weight of the i-th input corresponding to the j-th regular neuron, and $b_j(t)$ is the j-th deviation. In addition, n represents the number of input neurons in the input layer; M represents the number of regular neurons in the regular layer; and $f_j(t)$ and $f''_j(t)$ are the lower and upper bounds of the activation intensity of regular neurons, respectively. It can be expressed as follows:

$$
f'_j(t) = \sum_{i=0}^{n-1} u'_{ij}(t),
$$

$$
f''_j(t) = \sum_{i=0}^{n-1} u''_{ij}(t+1).
$$

3.2. Design of Self-Organization Mechanism of Self-Constructed Interval Type 2 Fuzzy Neural Network. In order to improve the performance of IT2FNN, the article proposes a self-constructed interval type 2 fuzzy neural network based on collaborative strategy, including a self-organization mechanism based on collaborative strategy and a parameter optimization mechanism based on collaborative strategy. The self-organization mechanism uses a comprehensive evaluation algorithm and a structure adjustment mechanism to make the structure adjustment of the self-constructed interval type 2 fuzzy neural network
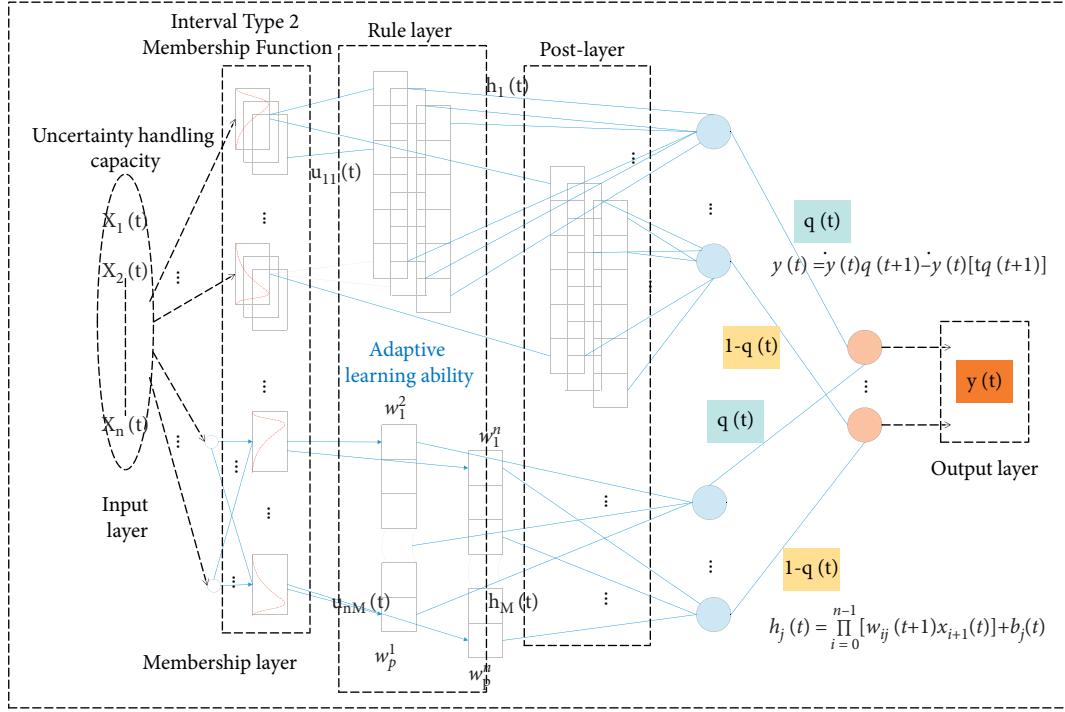
FIGURE 1: Interval type 2 fuzzy neural network structure.

coordinate with the parameter optimization process. The comprehensive evaluation algorithm uses the interneuron and interlayer information to comprehensively evaluate the structure of SC-IT2FNN and then uses the structure adjustment mechanism to add and delete fuzzy rules to realize the structural self-organization of SC-IT2FNN. In addition, SC-IT2FNN does not need to preset any threshold when making self-organizing judgments in the learning process. This feature is conducive to practical applications.

In IT2FNN, the activation intensity of regular neurons reflects the ability of fuzzy rules. Effective or redundant fuzzy rules can be found so that SC-IT2FNN has a suitable network structure. In the comprehensive evaluation algorithm, the similarity is as follows:

$$
\begin{aligned}
S_{ij}(t) &= \frac{\prod_{z=0}^{Z-1}\left[F_j(z-t-1)+F'(z+t+1)\right]\left[F_i(z+2t-1)-F''(z-t+2)\right]}{\left[F_j(z-t-1)-F'(z+t+1)^2\right]^{1/2} \cdot \left[F_i(z+2t-1)^2+F''(z-t+2)\right]^{1/2}}, \\[2mm]
F_j(t+z-1) &= 0.5\left[f_j'(t-2z-1)-f_j''(2t+z-1)\right], \\[2mm]
F''(t-2z-1) &= \prod_{j=0}^{M-1} F_j\left(\frac{t+2z}{M-1}\right), \\[2mm]
C_j(t) &= \frac{S_{ij}(t)}{d_j(t)}, \\[2mm]
d_j(t) &= \left\{\left[F_j(t+1)-Y(t-1)\right]V(t)^{-1}\left[F_j(t+1)-Y(t-1)\right]^{T}\right\}^{1/2}, \\[2mm]
F_j(t) &= \left[\,F_j(t-1)\ \ F_j(t)\ \ \dots\ \ F_j(t-1+Z)\,\right]^{-1}, \\[2mm]
Y_c(t) &= \left[\,y(t-1)\ \ y(t)\ \ \dots\ \ y(t-1+Z)\,\right]^{-1}.
\end{aligned}
\tag{3}
$$

The comprehensive evaluation algorithm uses the similarity and independent contribution of rule neurons to evaluate the effectiveness of fuzzy rules. Similarity can indicate the necessity of each fuzzy rule, and independent contribution degree can indicate the effectiveness of each fuzzy rule. Through these two evaluation indicators, it can be

judged whether each fuzzy rule in SC-IT2FNN is redundant and effective. At the same time, when calculating the similarity and independent contribution of fuzzy rules, the state value of the rule neuron is used to ensure the accuracy of the similarity and independent contribution.

In the design of this coordination mechanism, the connection of fuzzy weights and the independent transmission of degrees are three structural transformation states, which correspond to three stages of adjustment. Structure diagram of adjustment mechanism is shown in Figure 2.

3.3. Design of Parameter Optimization Mechanism for Self-Constructed Interval Type 2 Fuzzy Neural Network.

Traditional optimization algorithms such as backpropagation algorithm, gradient descent algorithm, Newton method, and Levenberg–Marquardt algorithm are widely used in the parameter optimization of neural networks, but these algorithms have the problems of slow convergence speed and difficulty in obtaining optimal solutions, such as backpropagation algorithm and gradient descent algorithm. The calculation process of Newton's method and Levenberg–Marquardt algorithm is complicated and time-consuming. There are many restrictions in the use process. Due to the numerous parameters of the interval type 2 fuzzy neural network, more calculation time is required in the process of optimizing the parameters.

In order to optimize the parameters of SC-IT2FNN, the article proposes a parameter optimization mechanism based on a collaborative strategy, that is, a hierarchical optimization mechanism. In this hierarchical optimization mechanism, the parameters are divided into nonlinear and linear parameters for collaborative optimization. Nonlinear parameters are updated using an improved second-order algorithm, and linear parameters are updated using a least-squares algorithm. This layered optimization mechanism can quickly update the SC-IT2FNN parameters, effectively improve learning accuracy, and reduce computational complexity. The hierarchical optimization mechanism includes two parts: parameter analysis and parameter optimization. In the parameter analysis, the SC-IT2FNN parameters including uncertain mean, standard deviation, scale factor, subsequent weights, and deviations will be divided into two types: nonlinear parameters and linear parameters. The following parameters are nonlinear parameters:

$$\Phi_N(t) = \begin{bmatrix} \sigma_{ij}(t-1) & q(t) & m_{ij}(t+1) \end{bmatrix}. \quad (4)$$

where $\Phi_N(t)$ is a nonlinear parameter set containing three nonlinear parameters. Therefore, SC-IT2FNN can be expressed as follows:

$$Y(t) = \Phi_L(t+1)E(t) + \Psi[x(t), \Phi_N(t)],$$
$$\Phi_L(t) = \begin{bmatrix} w_{ij}(t-1) & b_j(t) \end{bmatrix}. \quad (5)$$

There are two parts: updating nonlinear parameters using a modified second algorithm and optimizing linear parameters using at least two algorithms. But the usual parameter normalizer is the best. In addition, SC-IR2FNN

process is optimal for construction and general parameter modification, as well as nonlinear and general linear parameters. The improved quadratic algorithm and least square algorithm are used in the structure transformation process. The SC-IR2FNN network structure is used to adjust and total parameters. We downloaded it based on sales connect policy SC-IR2FNN and learned about the real stacking steps of processing. By designing a self-adjusting learning rate, the difficulty of choosing a learning rate is solved. For SC-IT2FNN, the layered optimization mechanism can effectively improve accuracy and reduce computational complexity. Using multiple sets of sample data to calculate the similarity and independent contribution of the ruled neurons not only avoids frequent adjustments to the network structure that may cause the network to fail to converge but also ensures the accuracy of the network structure evaluation and reduces the computational burden. Therefore, SC-IT2FNN has the characteristics of good generalization performance and fast calculation speed. The specific implementation process of the learning process of SC-IT2FNN is based on collaborative strategy as shown in Figure 3.

4. Result Analysis

4.1. Logistic Model and Data Selection.

In terms of supply chain financial risk management, foreign countries have developed a variety of methods to measure supply financial risk. Traditional methods include expert scoring method, fuzzy comprehensive evaluation method, credit rating method, BP neural network, and logistic model. The logistic model has the following advantages:

First, the logistic model has relatively simple requirements for data collection and processing and has strong operability; second, the model predicts the probability that the result is between 0 and 1, which can intuitively see the credit risk of the financing enterprise; Third, the model's preconditions are relatively loose and can be applied to continuous or categorical independent variables. Therefore, this paper finally chooses the logistic model to evaluate the credit risk of B2C e-commerce online supply chain finance.

Suppose the conditional probability of credit risk in financing enterprises is p. When the value of p is closer to 1, it indicates that the credit status of the enterprise is better; otherwise, the credit is not good. The logistic model does not theoretically have a critical value, and 0.5 is used as the critical value during model analysis. Therefore, this paper also uses 0.5 as the critical value when studying corporate credit risk. If the calculated value is greater than or equal to 0.5, the company's credit is considered good; otherwise, the company's credit is bad.

This article takes an e-commerce company's participation in online supply chain finance as an example, applies the evaluation index system established above, comprehensively evaluates the company's credit status based on the online supply chain financing model, and compares and analyzes the results. The example quantitatively analyzes the effectiveness of the logistic model on the financial credit risk control of the B2C e-commerce online supply chain.
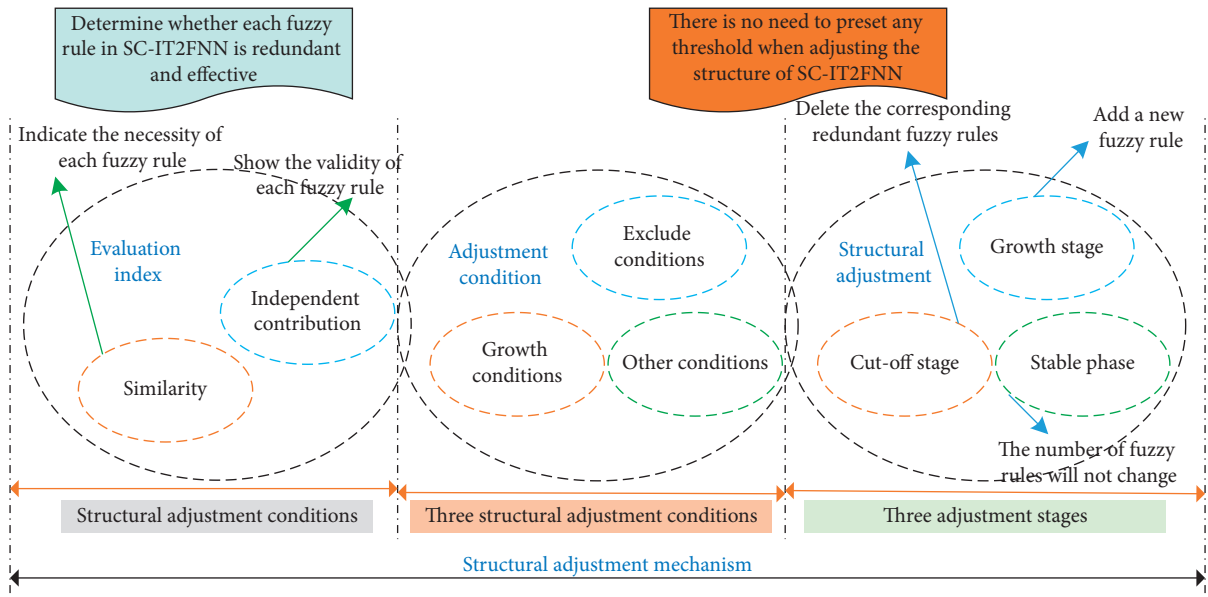
FIGURE 2: Structural adjustment mechanism.
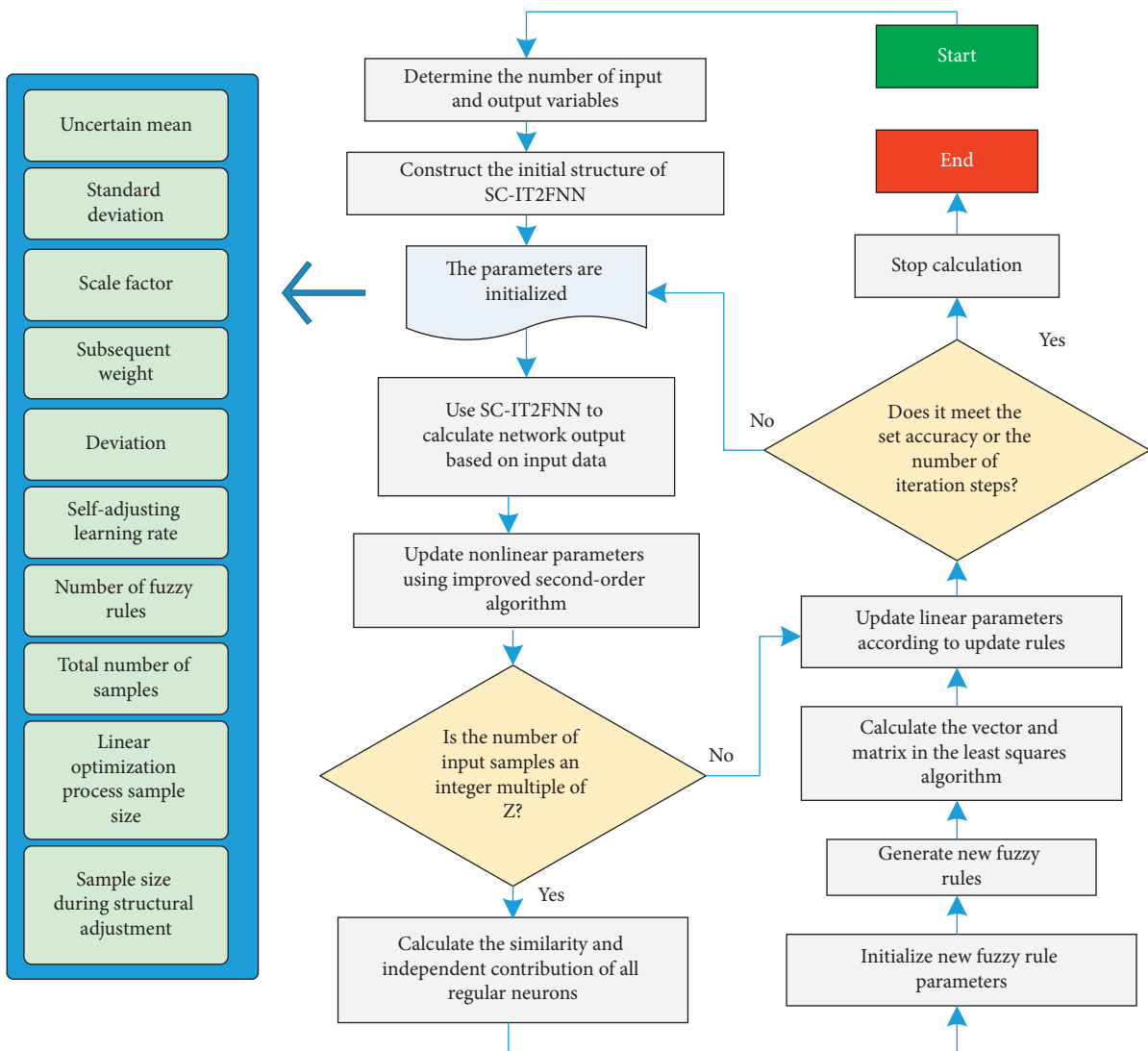
FIGURE 3: The specific implementation process of the learning process of SC-IT2FNN based on collaborative strategy.

E-commerce is one of the previous generation B2C online shopping platforms. There are many kinds of household appliances, such as traditional household appliances, 3C household appliances, and daily necessities. Online finance and e-commerce is one of the main directions. The supply of financial consumption platoon, investment asset management, tips, ointment, prepaid cards, preleasing, private equity financing, and other businesses have been launched. Its attempts in the internet field have further expanded the influence of an e-commerce company and realized collaboration and cooperation with upstream and downstream enterprises for a win-win situation.

The supply chain finance business of an e-commerce company is developed on the basis of its one-stop supply chain service platform. Customers can realize the transaction and financing process on the platform. However, as the business layout expands, an e-commerce company has also encountered applications. In order to help an e-commerce company better cope with the credit risk of small- and medium-sized enterprises in online supply chain finance, this article selected 40 e-commerce companies on an e-commerce platform and upstream and downstream companies in the supply chain from the Wonder database as samples. We take the lower value of the interest-bearing debt ratio of each industry in 2020 as the limit. Among the existing 40 sample companies, 6 companies have breached the contract, and 34 companies have not breached the contract.

4.2. Factor Analysis.

When using the logistic model to model the data, it is required that there should be no collinearity between the independent variables. Therefore, we first use factor analysis to standardize the variables and select representative independent variables to replace the original indicator variables. The newly acquired common factors are linear combinations of the original variables.

First, we need to check tapinle's data with Bartlett to see if the indicators selected by individuals are factor analysis. If KMO is 0.9 or greater, factor analysis is considered to be used. If it is 0.8–0.9, it is open. 0.6–0.8 spacing is normal. At one level, we then do factor analysis. The results of the Bartlett and KMO tests are shown in Figure 4.

The KMO test coefficient in this article is greater than 0.5, and the partial correlation between variables is strong. The Sig. of the Bartlett sphere test is less than 1%, indicating that we can do factor analysis.

4.3. Logistic Regression Model Analysis.

We have to substitute the logistic regression model for the seven principal components. The sample companies are used as dependent variables, and the seven principal component factors are used as independent variables. The factors F1~F7 are substituted into the logistic model for regression analysis using the entry method, that is, all variables are substituted into the equation at one time. The results of using SPSS software are shown in Figure 5.

It can be seen from the regression results that the final explanatory variables of the seven models are all retained,
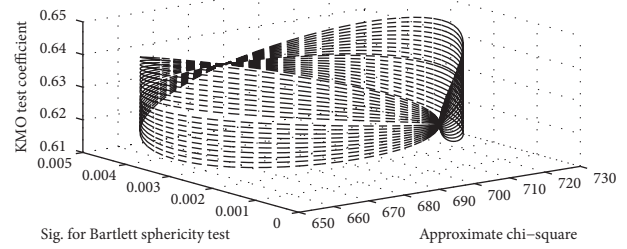
FIGURE 4: Bartlett and KMO tests.
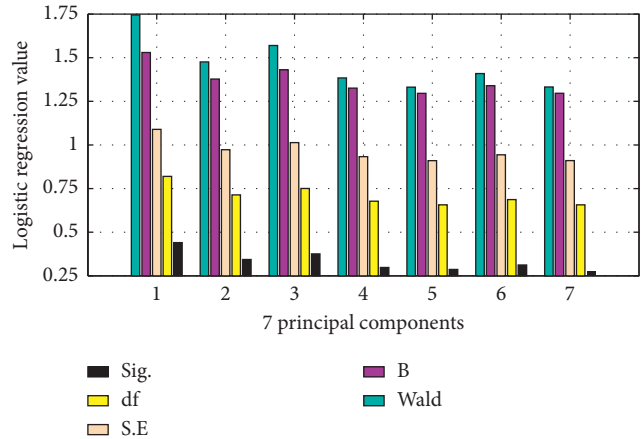
FIGURE 5: Logistic regression results.

and their Sig. values are all less than 0.5, indicating that these seven variables are very convincing for predicting the credit risk of financing enterprises.

4.4. Results and Verification of Model Analysis.

The seven principal component factors calculated by using SPSS software are used as independent variables of the model, and two types of enterprises (the value of enterprises with credit risk is 1, and the value of enterprises without credit risk is 2) are selected as dependent variables. The sample results are shown in Table 1.

As shown in Table 1, we can see that in these 40 samples, among the 29 risk-free companies that have been observed, 29 risk-free companies have been predicted using the model, with an accuracy rate of 100%. Among the 6 observed high-risk companies, 5 were predicted by the model to be risky companies, with an accuracy rate of 87%. Therefore, the final comprehensive accuracy rate of this model reached 89%. This shows that the prediction accuracy of the model is high. Table 2 shows the results of the significance test using SPSS.

As shown in Table 2, the Sig. value of the model test is 0.021, which is significantly lower than 0.05, indicating that the logistic regression equation obtained is significant at the 95% level. This illustrates that the seven factors, namely, the profitability factor of the financing enterprise, the solvency factor of the financing enterprise, the profitability factor of the core enterprise, the operation guarantee factor, the growth ability factor of the financing enterprise, the online degree factor of the supply chain, and the quality of the

TABLE 1: Classification of inspection samples

Observed		Predicted		
		E-commerce credit risk		%
		1	2	
E-commerce credit risk	1	6	0	87.01
	2	5	29	93.40
Overall percentage				89.25

TABLE 2: Comprehensive test of model coefficients.

Bangla	Sig.	df
14.2	0.021	6
13.9	0.021	7
15.1	0.021	5
15.2	0.021	7
14.7	0.021	6
13.8	0.021	5
15.4	0.021	7

financing enterprise and cooperation factors, have a significant relationship with the credit risk of the enterprise and further illustrates the practical application value of the model.

4.5. Anticycle Deception Analysis.

Here, we select a set of experimental data with this feature to compare and analyze the difference between the credit scoring method proposed in this article and Taobao's credit scoring method. The experimental results are shown in Figure 6. It can be seen from Figure 6 that the scoring method using RBF neural network credit accumulation is completely powerless against periodic deception. The credit scoring method of a fuzzy neural network describes the process of cyclical deception in a very specific way and inhibits the growth of credit value, which is conducive to the risk judgment of buyers and can also play a supervisory role for sellers. Within the transaction failure threshold, the credit value will not be degraded, and when the threshold is exceeded, the credit value will be degraded. Therefore, the setting of the threshold of transaction failure rate is critical. If the number of negative reviews specified by the transaction failure rate threshold is not reached, the credit value begins to degrade, indicating that the threshold is set too large and the actual transaction failure rate is small. Conversely, if the number of times specified by the threshold is reached and the credit value has not been degraded, it means that the threshold setting is too small, and the actual failure rate is relatively large.

4.6. Analysis of Anticredit Hype.

We select realistic data that meet this situation for analysis, and the experimental results are shown in Figure 7.

Under Taobao's credit scoring method, although the seller did not cheat after the credit speculation, the credit at this time has been greatly reduced, and the buyer is easily confused by the credit value. It can be seen from Figure 7
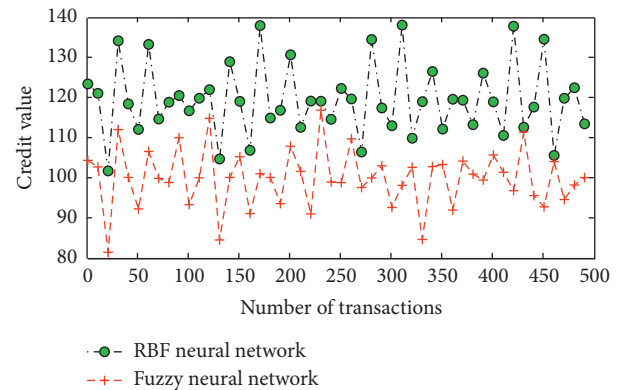
FIGURE 6: Comparison of the credit value growth process of cyclic deception.

(a)

(b)
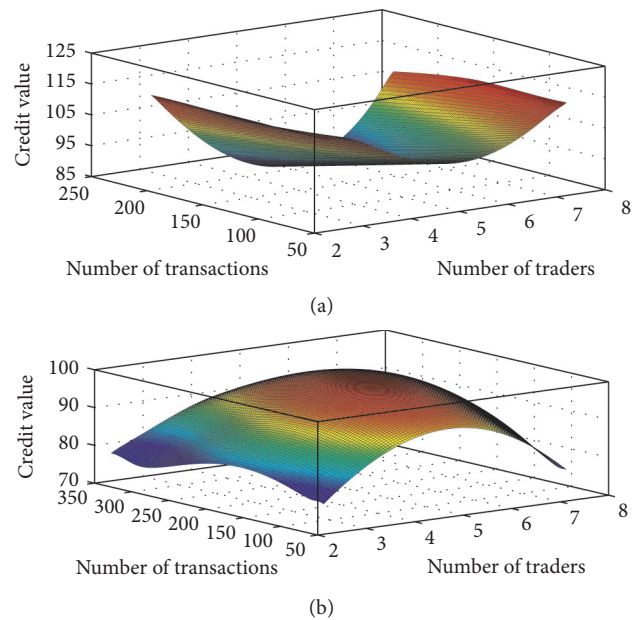
FIGURE 7: The credit value of the cycle deception after the credit hype: (a) results of RBF neural network method and (b) the results of the fuzzy neural network method.

that the credit scoring method of the RBF neural network cannot curb fraudulent behavior after credit speculation. And through the fuzzy neural network, it can be clearly seen that the deceptive credit value after the credit hype has been reduced.

5. Conclusion

From the empirical results, it can be seen that indicators such as net sales interest rate, return on net assets, and net profit rate of total assets are positively correlated with the profitability factor of financing enterprises. At the same time, the profitability factor of financing enterprises is positively correlated with the compliance probability of online supply chain finance SMEs. Therefore, the higher the net profit margin of the financing company's sales and the net profit margin of the financing company's total assets, the lower the credit risk of SMEs participating in online supply chain financing. The solvency of financing companies is negatively related to the probability of online credit risk. That is, the stronger the solvency of SMEs, the smaller the credit risk in their online supply chain financing business. From the empirical results, we can see that the current ratio, quick ratio, and interest coverage are positively correlated with solvency, and the debt-to-asset ratio is negatively correlated with solvency. Therefore, the higher the current ratio, quick ratio, and interest protection multiple, the stronger the solvency of SMEs, and the lower the credit risk of their participation in online supply chain financing; the higher the debt-to-asset ratio, the lower their solvency. The online degree factor is negatively related to the probability of the occurrence of financial credit risk in the online supply chain. That is to say, the stronger the online degree of this business, the smaller the credit risk of this online supply chain financing business. It can be seen from the component score coefficient matrix that the indicators reflecting the online degree of the supply chain, such as the electronic order processing capacity and the degree of information sharing, are positively correlated with the online degree factor of supply chain finance, and the online degree factor is related to the compliance of small- and medium-sized enterprises. The probability is positively correlated. Therefore, the higher the electronic order processing capability and the degree of information sharing, the lower the credit risk of SMEs participating in online financing business. When dealing with the problem of using low-priced commodities for credit speculation, the improved credit scoring method considers prices in different segments, and the increase in credit value is related to the ratio of price segments in all price segments. The closer the price range of the goods sold, the faster the credit growth of sellers. Once the price range of the goods sold by the seller fluctuates sharply, the growth of the credit value slows down. In this way, credit hype can be curbed while protecting the interests of sellers who only sell low-priced goods. Comparative experiments show that the self-constructed interval type 2 fuzzy neural network is effective in suppressing credit speculation. When dealing with the problem of periodic fraud, the improved credit scoring method sets different credit levels according to the level of credit value. In different credit levels, different deduction coefficients are set. A high credit rating sets a high deduction factor to increase the opportunity cost of periodic deception. Through this kind of punishment, the cycle of deception is effectively curbed. The comparative experiment shows that the self-constructed interval type 2 fuzzy neural network is effective in punishing the periodic deception.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] C. Wei, Z. J. Yu, and X. N. Chen, "Research on social e-commerce reputation formation and state-introduced model," *Kybernetes*, vol. 46, no. 6, pp. 1021–1038, 2017.

[2] P. J. Wu and K. C. Lin, "Unstructured big data analytics for retrieving e-commerce logistics knowledge," *Telematics and Informatics*, vol. 35, no. 1, pp. 237–244, 2018.

[3] Y. Zhang, "Building China EU cross border e-commerce ecosystem with blockchain technology," *China's Circulation Economy*, vol. 2, no. 2, pp. 66–72, 2018.

[4] Y. Hao, M. Usama, J. Yang, M. S. Hossain, Q. Liu, and A. Ghoneim, "Recurrent convolutional neural network based multimodal disease risk prediction," *Future Generation Computer Systems*, vol. 92, no. 1, pp. 76–83, 2019.

[5] P. Sulikowski, T. Zdziebko, D. Turzyński, and E. Kańtoch, "Human-website interaction monitoring in recommender systems," *Procedia Computer Science*, vol. 126, pp. 1587–1596, 2018.

[6] N. Dong and X. Zhu, "Blockchain technology evolution and industrial application prospect," *Information Security Research*, no. 3, pp. 200–210, 2017.

[7] S. Mohammed, "Bankruptcy prediction by using the altman Z-score model in Oman: a case study of raysut cement company SAOG and its subsidiaries," *Australasian Accounting Business and Finance Journal*, vol. 10, no. 4, pp. 70–80, 2016.

[8] C. Fan, J. Wang, W. Gang, and S. Li, "Assessment of deep recurrent neural network-based strategies for short-term building energy predictions," *Applied Energy*, vol. 236, pp. 700–710, 2019.

[9] Q. Wu, Q. Zhu, and P. Li, "A neural network based reputation bootstrapping approach for service selection," *Enterprise Information Systems*, vol. 9, no. 7, pp. 768–784, 2015.

[10] K. Braunlich, C. A. Seger, K. G. Jentink, I. Buard, B. M. Kluger, and M. H. Thaut, "Rhythmic auditory cues shape neural network recruitment in parkinson's disease during repetitive motor behavior," *European Journal of Neuroscience*, vol. 49, no. 6, pp. 849–858, 2019.

[11] H. Li, "Using blockchain technology to promote the development of cross border e-commerce in China," *Monthly Finance and Accounting*, no. 3, pp. 66–146, 2019.

[12] Z. F. Liu, "Analysis on the construction of cross-border E-commerce credit evaluation system under the background of big data," *China Business and Market*, vol. 30, no. 6, pp. 58–64, 2016.

[13] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," *Expert Systems with Applications*, vol. 128, pp. 84–95, 2019.

[14] O. Tibermacine, C. Tibermacine, and F. Cherif, "Estimating the reputation of newcomer web services using a regression-based method," *Journal of Systems and Software*, vol. 145, pp. 112–124, 2018.

[15] K. Bortko, P. Bartków, J. Jankowski, D. Kuras, and P. Sulikowski, "Multi-criteria evaluation of Recommending Interfaces towards habituation reduction and limited negative impact on user experience," *Procedia Computer Science*, vol. 159, pp. 2240–2248, 2019.

[16] Z. Lu and X. Ge, "Research on the application of blockchain in cross border payment," *Southwest Finance*, no. 2, pp. 23–28, 2018.

[17] T. Zhou, K. H. Thung, X. Zhu, and D. Shen, "Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis," *Human Brain Mapping*, vol. 40, no. 3, pp. 1001–1016, 2019.

[18] B. Yao, H. Hagras, M. J. Alhaddad, and D. Alghazzawi, "A fuzzy logic-based system for the automation of human behavior recognition using machine vision in intelligent environments," *Soft Computing*, vol. 19, no. 2, pp. 499–506, 2015.

[19] B. Ahirwadkar and S. N. Deshmukh, "Deep neural networks for recommender systems," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 12, pp. 4838–4842, 2019.

[20] Y. Tian, G. Zhao, and L. Shen, "Blockchain transportation: a case study of freight logistics and its market governance," *China's Circulation Economy*, no. 2, pp. 50–56, 2018.

[21] W. Zhu and G. C. Beroza, "PhaseNet: a deep-neural-network-based seismic arrival-time picking method," *Geophysical Journal International*, vol. 216, no. 1, pp. 261–273, 2019.

[22] A. Goswami, G. S. Parashari, and R. Gupta, "Evolutionary stability of reputation-based incentive mechanisms in P2P systems," *IEEE Communication Letter*, vol. 22, no. 2, pp. 268–271, 2017.

[23] S. Akuma, R. Iqbal, C. Jayne, and F. Doctor, "Comparative analysis of relevance feedback methods based on two user studies," *Computers in Human Behavior*, vol. 60, pp. 138–146, 2016.

[24] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, and W. Geng, "A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface," *Pattern Recognition Letters*, vol. 119, pp. 131–138, 2019.

[25] T. Oliveira, M. Alhinho, P. Rita, and G. Dhillon, "Modelling and testing consumer trust dimensions in e-commerce," *Computers in Human Behavior*, vol. 71, pp. 153–164, 2017.

[26] S. Rouhani, A. Ashrafi, A. Zare Ravasan, and S. Afshari, "The impact model of business intelligence on decision support and organizational benefits," *Journal of Enterprise Information Management*, vol. 29, no. 1, pp. 19–50, 2016.

[27] S. Lodhi, M. A. Manzar, and M. A. Z. Raja, "Fractional neural network models for nonlinear riccati systems," *Neural Computing and Applications*, vol. 31, no. 1, pp. 359–378, 2019.

[28] H. Fang, L. Xu, and X. Huang, "Self-adaptive trust management based on game theory in fuzzy large-scale networks," *Soft Computing*, vol. 21, no. 4, pp. 907–921, 2017.

[29] A. Singraber, J. Behler, and C. Dellago, "Library-based LAMMPS Implementation of high-dimensional neural network potentials," *Journal of Chemical Theory and Computation*, vol. 15, no. 3, pp. 1827–1840, 2019.

[30] B. Li, Y. Li, H. Wang, Y. Ma, Q. Hu, and F. Ge, "Compensation of automatic weighing error of belt weigher based on BP neural network," *Measurement*, vol. 129, pp. 625–632, 2018.

Research Article

Logic Design and Power Optimization of Floating-Point Multipliers

Na Bai, Hang Li, Jiming Lv, Shuai Yang, and Yaohua Xu [ID]

Key Laboratory of Computational Intelligence and Signal Processing, Ministry of Education (Anhui University), School of Integrated Circuit, Anhui University, School of Electronic Information Engineering, Anhui University, Hefei 230601, China

Correspondence should be addressed to Yaohua Xu; xyh@ahu.edu.cn

Under IEEE-754 standard, for the current situation of excessive time and power consumption of multiplication operations in single-precision floating-point operations, the expanded boothwallace algorithm is used, and the partial product caused by booth coding is rounded and predicted with the symbolic expansion idea, and the partial product caused by single-precision floating-point multiplication and the accumulation of partial products are optimized, and the flowing water is used to improve the throughput. Based on this, a series of verification and synthesis simulations are performed using the SMIC-7 nm standard cell process. It is verified that the new single-precision floating-point multiplier can achieve a smaller power share compared to the conventional single-precision floating-point multiplier.

1. Introduction

According to a study by Stanford University, the demand for arithmetic power in artificial intelligence has doubled every three to four months since 2012, a rate that has surpassed Moore's law (doubling the number of transistors in a chip every 18 months) [1]. That is, the current rate of Moore's law has lagged far behind the speed of the growth in demand for computing power, relying solely on the progress of process technology that has been entirely unable to meet the demand for computing power growth. According to Intel's latest study, for computationally intensive applications, floating-point operations account for 75% of processor core power consumption and 45% of total processor power consumption [2], as shown in Figure 1, while its operator power consumption accounts for 39.57% in Think-II, an artificial intelligence inference chip developed by the first team of the Institute of Microelectronics at Tsinghua University [3]. From this, it can be seen that floating-point operation is a module with a very large power consumption ratio in big data computing, so it is significant to reduce the power consumption of floating-point operation.

As an important indicator of scientific computing capability, floating-point arithmetic is an important computing method for processing various information in today's electronic information tools. Floating-point numbers have a large dynamic range compared to fixed-point numbers; however, the structure is more complex and power consumption is higher, so it is a meaningful breakthrough to achieve multiplicative power optimization of floating-point numbers. In the past, 60% of the performance improvement of microprocessors depended on process advancement [4]; now, we need a breakthrough in design to achieve the demand of arithmetic power growth.

This paper focuses on the power optimization of single-precision floating-point multiplication operations in floating-point operations. The single-precision floating-point multiplication calculation can be divided into three parts, namely, the judgment of symbolic bits, the summation of exponential bits, and the multiplication of trailing bits. The first two conventional designs are similar in this paper, but the core of the design optimization is in the multiplication of trailing bits.
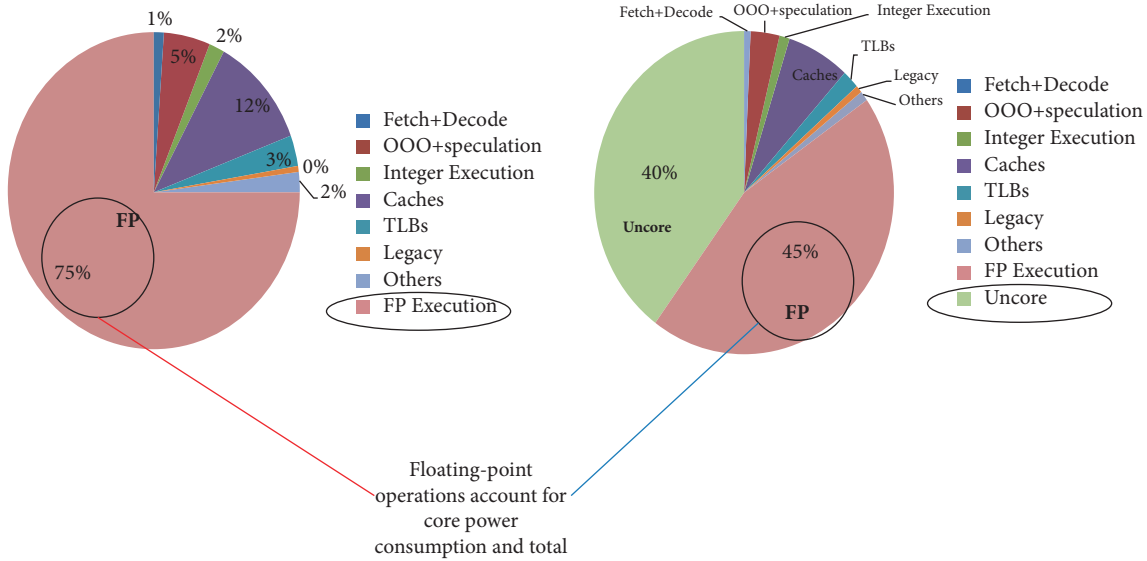
Figure 1: Intel's microprocessor core-level.

The conventional design for the trailing part of the multiplier is mainly a Wallace tree structure with a combined CSA and 4-2 compressor, mainly in the form of a booth code. Finally, the final product is obtained with an overrunning adder [5]. Figure 2 shows the conventional tail partial multiplier design.

In this paper, we hope to improve the coherent design to reduce the power consumption of single-precision floating-point multipliers.

2. Materials and Methods

2.1. Ideas and Advantages of the Design Algorithm Based on Boothwallace Tree. For the current design of single-precision floating-point multiplication, this paper divides the design into three parts as the direction of optimization: how to appropriately increase the throughput of data processing when emitting signals; how to process the partial products caused by the multiplication of trailing bits and reduce the number of operations and storage at the same time; and how to effectively reduce the workload of summing and reduce the hyperactive operations in the process of partial product summing.

This design is considered for processing a huge amount of multiplication data operations at 2 GHZ frequency. If we just follow the timing sequence to process one data and then continue to process the next data, such a way is actually very inefficient. Therefore, this paper adopts the form of flowing water to improve the throughput of data processing and increase the frequency of the clock. Although the primary delay $(n * (T1 + T2))$ is added in the first stage of processing, the pipelined operation can substantially improve the overall data processing efficiency because the register delay must be smaller than the combinational logic delay. The difference time between the register delay and the combinational logic delay can be saved almost every time the pipelined data processing is done under the huge data processing. The

disadvantage of pipelining is that it increases power consumption, area, and hardware complexity. Therefore, while using pipelining to improve data processing efficiency, it is important to weigh the loss of power consumption, area, and software complexity.

The encoding operation of the partial product caused by the trailing bits in the multiplication stage can significantly affect the final power consumption and area. If the partial product is used directly or a simple booth-2 coding operation is done, the power consumption loss and the area occupied are too large. Adopting the booth-4 encoding form can reduce half of the partial product caused by booth-2, compared to the direct use of the partial product form that reduces the part beyond 1/2, which can significantly reduce the time consumption caused by the subsequent operation of partial product, but also can significantly reduce power consumption. Although in the case of booth-4 coding operation, there will be some assignment and logic operations, which will lead to this coding operation will consume a certain amount of time and power consumption, as well as occupy part of the area, still, for a large number of data operations, the increased consumption of this part is much smaller than the consumption saved by the subsequent partial product operation of the data, so this paper uses the booth-4. The symbolic bit expansion method is used in this paper [6]. The symbolic bit expansion can predict the partial product sum form, which can appropriately compress some redundant operations of the partial product sum, and use a certain logical relationship between the two summed partial products to establish the variable relationship, which can compress the area and reduce the power consumption to a certain extent to achieve a comprehensive multiplication of multiple bits, and a large number of partial sums can save a lot of areas and power consumption effect and for subsequent operations can reduce more add operations [7].

In the partial product summation stage, using the direct array summation form or the rounding reservation form will
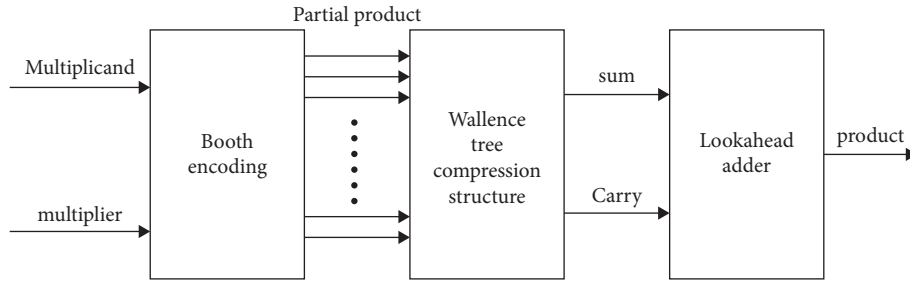
Figure 2: Conventional multiplier design.

result in a large blanking hypervalue, which will not only occupy a large area but also increase the nonessential addition operations leading to increased power consumption and time. Using the Wallace tree in the form of overlay summation in the form of tree depth reduction will greatly reduce the number of nonessential addition operations and reduce the amount of white space, resulting in a significant improvement in time. Although multiple uses of the full adder can cause a slight loss of power consumption, but for a large amount of data processing, it brings better rate gain and is therefore chosen [8].

2.2. Implementation Platform and Implementation Algorithm

2.2.1. Implementation Platform and Process. Figure 3 shows the experimental platform and the validation process [9].

The first step is to build a golden model database in the MATLAB platform about single-precision floating-point multipliers and multiplied numbers and generating control group products. Since the range of single-precision floating-point designs is too wide to be verified by the exhaustive method based on hardware limitations, it is the idea of the design to create a database of golden models as the basis for verification calls. The simple steps are to use the rand function and the num2hex function to generate random data and convert single-precision floating-point data. A database of 100 million multipliers and multiplied numbers are created, saved as num and hex types, respectively, and multiplied together in MATLAB to form a product database as a control group for later testing processes as an accuracy standard. After generating the data, the hist function is used to demonstrate the uniformity of the generated database to ensure that the generated database is the golden model database and reduce the uncontrollable experimental error. The following are the functions for generating random data and converting type functions:

$$a(i) = f + 2 * e * \text{rand}(1, 1, \text{single}),$$
$$a\_\text{temp} = \text{num2hex}(a(i)). \tag{1}$$

Under the foundation of the golden model database with good calls, a suitable testbench environment needs to be established as a way to verify the authenticity of the source code logic and debug the source code logic optimization. The platform for design implementation is the NC-Verilog platform. Here, in addition to calling the golden model database and generating the verification

product library, the direct window output under the waveform file is also done. The display function is called with embedded precision verification judgments to simplify verification access.

In order to subsequently implement the power consumption parameter analysis and obtain the simulation data for each of the source codes, a synthesizable implementation of the source code is required. The design is implemented in the Design Compiler platform. The original description of the source code is modified, and the source code is reshaped in the specified hardware description language. After invoking the synthesis environment, the synthesis simulation of the source code is realized, and the simulation data of each source code can be obtained, and the required slack data and area data can be extracted as the final control index. In the process of synthesis, the required netlist file is generated as a sample for subsequent calls. It should be noted that the DC environment CLK needs to be annotated when calling Design Ware samples, and the source code constraint in this paper is 2 GHz. Here, this paper relies on the final code to generate the circuit netlist to observe the structure of the circuit diagram.

After generating the required circuit netlist file samples, call the samples to verify them in the testbench environment, noting that it may be necessary to create a process library for import. After verification, the VCD waveform file sample is exported, and the circuit netlist sample and VCD waveform file sample are imported into the PTPX platform to create the required PTPX environment to generate the corresponding power analysis report sample. Based on the samples and the required theory, the source code is designed and optimized, and the optimized samples are saved. And save the Design Ware samples as a control group to achieve the desired results.

2.2.2. Logic Design and Optimization. The basic single-precision floating-point multiplier is built on the basis of the binary floating-point representation formula, which extracts the sign bits, the tail f, and the step code e from the 32-bit binary data, respectively. Figure 4 represents the arithmetic structure of this 32-bit single-precision floating-point number. The core part is the fixed-point multiplication of the trailing part af:{1′b1, a[22:0]}, bf:{1′b1, b[22:0]}. In the based code, this multiplication structure is the simplest shift judgment flow structure. In the optimization structure, the main thing is also the fixed-point multiplication in the tail
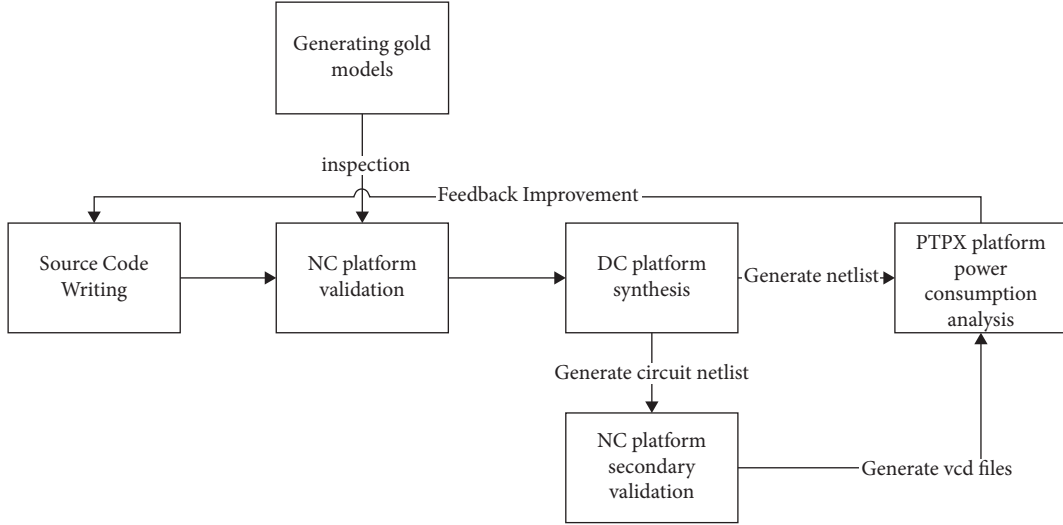
FIGURE 3: Design verification and simulation flow.
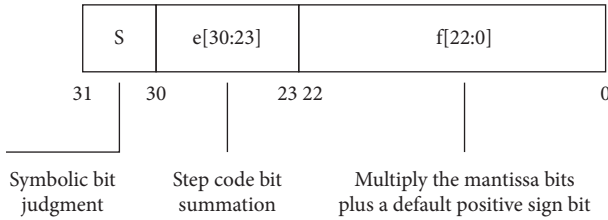
FIGURE 4: Single-precision floating-point multiplication structure.

TABLE 1: Booth-4 transcoding table.

$X_{2i+1}X_{2i}X_{2i-1}$	Part of the product $PP_{[n-1:1]}$	NEG	Z	B_1	B_2
000	$+0 * Y$	0	1	1	0
001	$+1 * Y$	0	1	0	1
010	$+1 * Y$	0	0	0	1
011	$+2 * Y$	0	0	1	0
100	$-2 * Y$	1	0	1	0
101	$-1 * Y$	1	0	0	1
110	$-1 * Y$	1	1	0	1
111	$-0 * Y$	1	1	1	0

part and the optimization of the accuracy in the combination of the final product c parts [10].

The booth-4 encoding has been introduced above to effectively reduce the number of partial products caused by the fixed-point multiplier. In this paper, the single-precision floating-point multiplier is actually optimized by the booth-4 encoding for the trailing part (i.e., the 24-bit fixed-point multiplier). In addition, this paper also deals with the symbolic bit expansion of the booth-4 encoding. Since this 24-bit fixed-point multiplier is treated as a symbolic number by default, a simple processing should be done. af is processed as aff:{1′b0, af}. In this paper, bf is treated as the transcoded original code of booth-4 in the code as bf_b: {2′b00, bf, 1′b0} form, and then the transcoding process is performed corresponding to Table 1, hosting the symbols of each part (neg) and generating the transcoding code (zero, one, two). According to the transcode, the aff is operated accordingly, with zero being taken, one being unchanged, and two being shifted left by one. Then, determine whether the operation of taking the inverse plus one is needed according to the neg symbol. There is an important logic here such that the neg symbol is 0 by default under the premise of determining zero. This is an important logic point in the subsequent symbolic bit expansion [11].

In the case that the basic encoding has been completed, it is necessary to do another simplification of booth-4, which is used as a prediction of the booth algorithm with a rounding, again compressing its partial product [12]. Then, according

to the partial product logical association of each line to do symbolic bit expansion simplification, here note that if there are zero cases, the front added symbolic bit neg default is 0, ~neg is 1. So far, we get 13 partial product prod, the end of the first flow level, into the second level of flow. Figure 5 shows the comparison of the partial product after the expansion with symbolic bits after the booth processing.

In the second level of flow, what this paper has to do is based on the first level of the partial product of 13 prods' multiple coverage compression to the last only one layer of product situation. According to the 3-2 compressor theory (2-2 compression is 3-2 compression in one additive to zero states), make add compression module; in the second level of flow in six levels cover the compression of 13 layers of prod so as to get the final result product. Therefore, a compression module needs to be added at each level to avoid bias. In the last level of the six-level overlay compression, the last two levels are added directly to obtain the final fixed-point 24-bit multiplication product using the system's built-in overfeed adder [13]. Figure 6 shows the compression coverage of the six-level level.

In the Wallace tree above, we have obtained the product of fixed-point multiplication results for the trailing part so that the symbolic bits, the order part, and the trailing part of the resultant floating-point number have been obtained. Nowadays, rounding and direct rounding are popular in
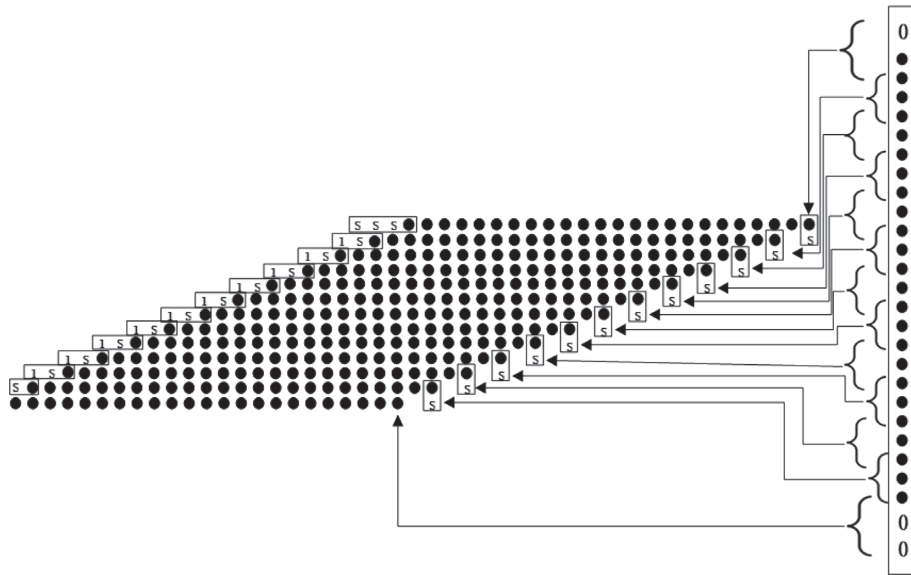
Figure 5: Schematic diagram of symbol bit expansion.
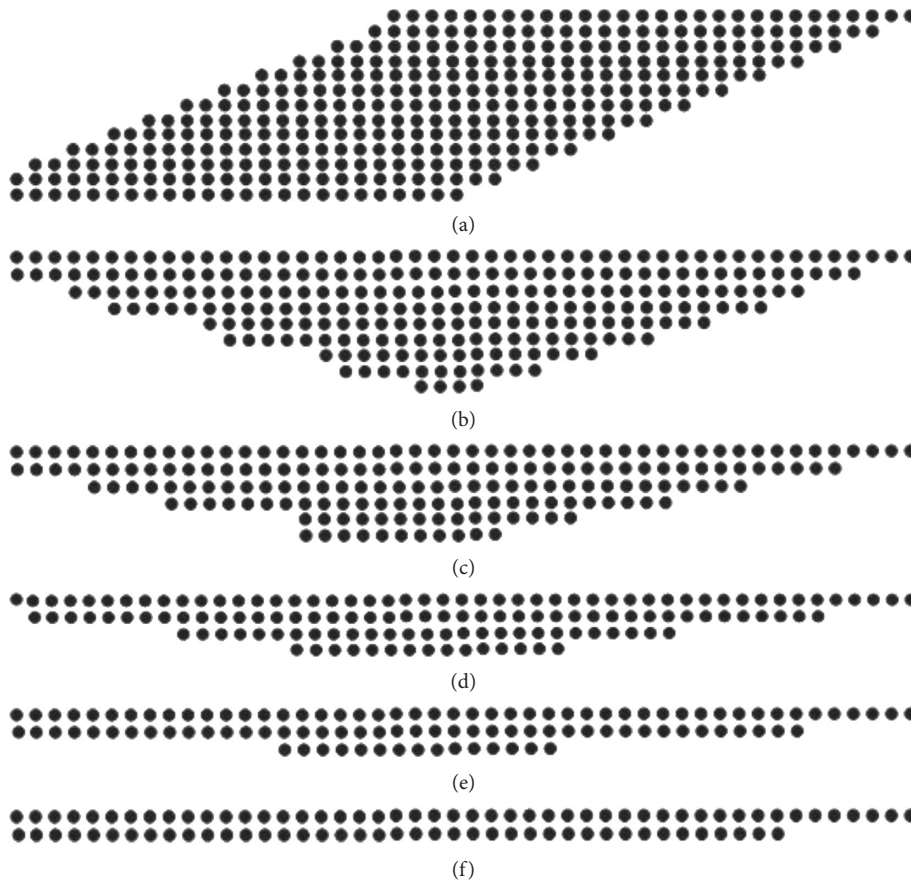
Figure 6: (a) The first-level compression diagram. (b) The second-level compression diagram. (c) The third-level compression diagram. (d) The fourth-level compression diagram. (e) The fifth-level compression diagram. (f) The sixth-level direct overfeed summation.

precision, and the former scheme is adopted in this design. In the end, the first result of the final fixed-point product is judged to be 1 to determine the overflow of the ordinal code. If the first place of fixed-point product is 0, the order code exceeds 382 as the maximum overflow and is below 127 as the minimum overflow, and the fixed-point product result cf2<=[45:23] is extracted in the nonoverflow case. If the first place of the fixed-point product is 1, the order code beyond
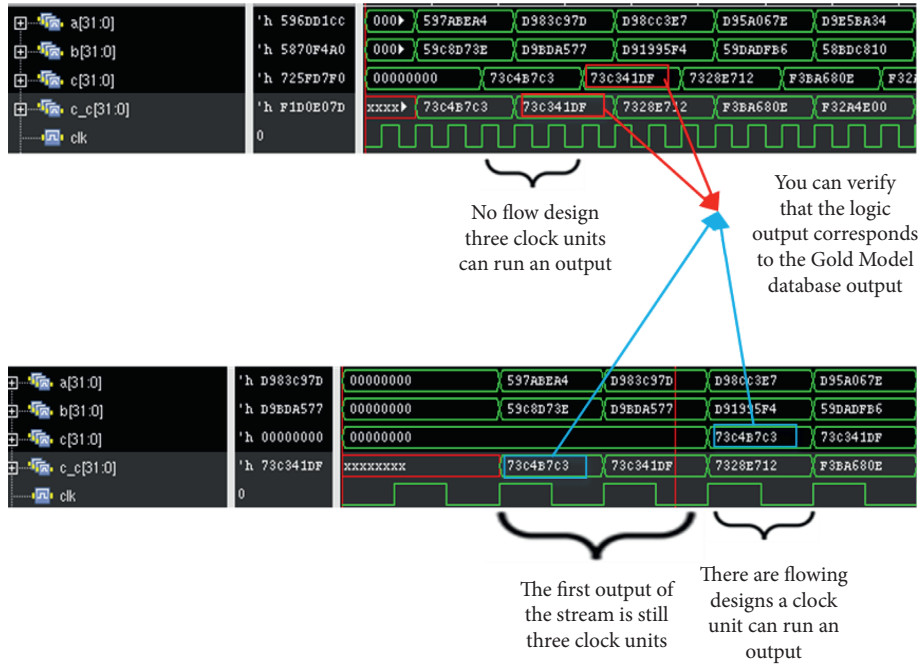
Figure 7: Logic output diagram of flowing and nonflowing water.
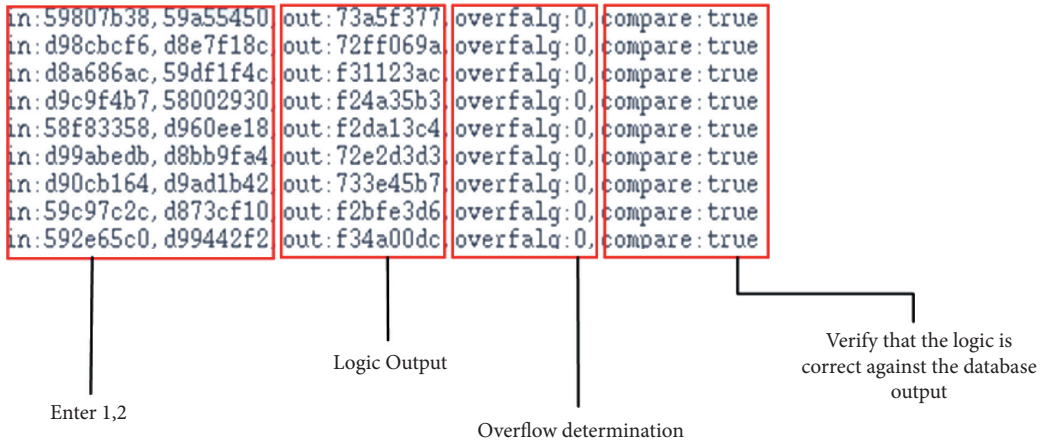
Figure 8: Logic function verification diagram.

381 is the maximum overflow, below 126 is the minimum overflow, and the result of extracting the fixed-point product in the nonoverflow case is cf5<=[46:24]. Then, do the precision constraint part, judge the latter bit of the extracted result part, and judge whether it is required as a feed. Here, it should be noted that the condition needs to be attached whether the extracted part is all one, or the order code is the maximum number 255, that is, whether the overflow result may be caused in the rounding state, if not, then the extracted part cf5 + 1. So, this paper gets a single-precision floating-point multiplier with the accuracy to meet the usage requirements. Complete the design requirements.

3. Results and Discussion

3.1. Code Logic Functional Verification

3.1.1. Testbench Environment Creation. Here, in this paper, we first write the most basic single-precision floating-point multiplier, whose core 24-bit fixed-point multiplier uses the simplest shift-hosted binary multiplication, and here we can get a series of codes for single-precision floating-point multiplier. After writing the codes, we make the corresponding comparisons on the flowing water, citing the nonflowing water format and the flowing water format. And
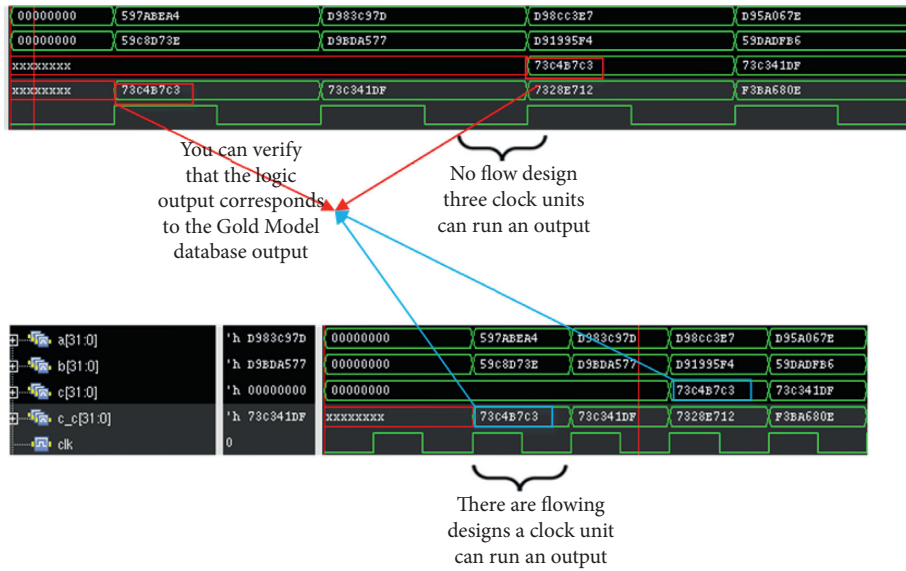
FIGURE 9: Optimized logic output diagram for flowing and nonflowing water.


```
in:594e5510,59e77886,out:73ba8fe5,overfalg:0,compare:true
in:597efc4c,598371c6,out:7382ec6d,overfalg:0,compare:true
in:582c65a0,d95f2aea,out:f216496d,overfalg:0,compare:true
in:d9829fc4,d95acf66,out:735f4bc0,overfalg:0,compare:true
in:59ed1ff4,d989e24c,out:f3ff6f6c,overfalg:0,compare:true
```
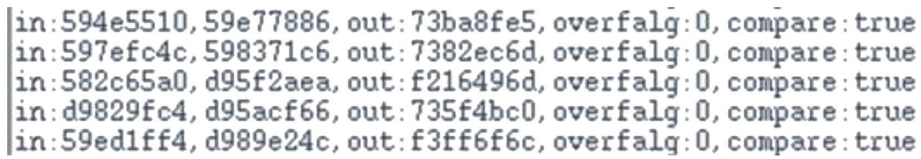
FIGURE 10: Final code logic verification diagram.

the source code is verified by the relevant logic, and the corresponding verification code is written in testbench, and the comparison output true for the code output is consistent with the MATLAB gold model production product result. See waveform in Figure 7 and logic verification in Figure 8.

As you can see, after the source code is written, the logic function is verified to be error free, and the accuracy is fully up to the requirements. The comparison between the waveform graphs of nonflowing and flowing formats also shows that the throughput in the flowing format is much larger than that in the nonflowing format so that the basic code is established, and the following is the code modification of the source code, adding the booth and Wallace tree modules, after which the combination and generation of partial product modules are optimized accordingly to achieve the desired goal.

3.1.2. Logic Validation after Adding the Boothwallace Tree Module.
In this paper, we add the optimized boothwallace module here to generate the final code, and then add and reduce the state of the running module, and use the original testbed environment to functionally verify the final code, and again compare the raw product of the code with the raw product in the MATLAB gold model database for numerical comparison. Figures 9 and 10 show the waveforms obtained from the verification.

Here, the verification related to the logic function of the code has been concluded in this paper, and the algorithm is

extracted and analyzed for each parameter to verify whether it achieves the expected effect of this paper when the use of the function can be confirmed without errors.

3.2. Comparison of Power Consumption of Fixed-Point Multipliers.
Here, in this paper, we first analyze the power consumption of the fixed-point multiplier in this paper and take out the 24-bit fixed-point multiplier of the most basic fmpy shift-hosting direct sum, the multiplier of the simple added booth module and Wallace module, and finally, the 24-bit fixed-point multiplier of boothwallace after optimization in this paper. Because these are all combined logic parts, this paper sets its max_delay to 0.5 ps, that is, according to the frequency of 2 GHz work, combines logic area, and measures its power consumption data. The data are recorded and compared at 25c, 85c, and 125c for three temperature process angles to get a relatively detailed report. They are displayed as Tables 2, 3, and 4, respectively.

The obtained statements are organized to obtain area (Figure 11) and power consumption comparison (Figure 12) for the three fixed-point module algorithms corresponding to different temperature process angles.

Here, we get the corresponding power consumption data of each fixed-point multiplier module. This paper can analyze that, with the maximum delay parameter set, the area of the combinational logic part of the original fmpy_24 fixed-point multiplier is the smallest under the process angle model of each temperature, resulting in its power

TABLE 2: Power consumption table of each fixed-point multiplier of 25c.

2 GHz (TC-25C)	max_delay	Area (combination)	Power
fmpy_24	0.5	298.752	$4.159E-06$
normal_24	0.5	529.4208	$2.934E-04$
boothwallace_24	0.5	324.2592	$1.923E-04$

TABLE 3: Power consumption table of each fixed-point multiplier of 85c.

2 GHz (TCH-85C)	max_delay	Area (combination)	Power
fmpy_24	0.5	298.752	$2.894E-05$
normal_24	0.5	529.4208	$2.947E-04$
boothwallace_24	0.5	324.2592	$2.116E-04$

TABLE 4: Power consumption of each fixed-point multiplier of 125c.

2 GHz (WC-125C)	max_delay	Area (combination)	Power
fmpy_24	0.5	298.752	$2.345E-05$
normal_24	0.5	529.4208	$3.397E-04$
boothwallace_24	0.5	324.2592	$2.309E-04$
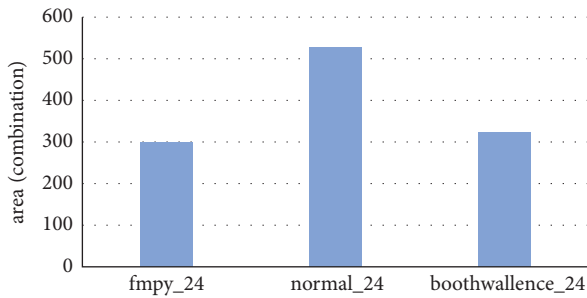
FIGURE 11: Comparison of the combined area of fixed-point modules.

FIGURE 12: Comparison of total power consumption of fixed-point modules at three temperature process angles.

consumption in all aspects being lower than the normal value. The area of the fmpy_24 fixed-point multiplier is the smallest, resulting in a lower than normal power consumption in all aspects. The area of the fmpy_24 fixed-point multiplier is increased by simply adding the booth code and the Wallace tree module, and the power consumption is increased accordingly. Since the fixed-point multiplication part is the core operation module of single-precision floating-point multiplication, the optimized power consumption and area of the fixed-point multiplication module are of great significance for the power consumption and area of the entire single-precision floating-point part. The following is a comparison of the power consumption of the single-precision floating-point section.

3.3. Single-Precision Floating-Point Multiplier Power Consumption Comparison. After the analysis of each bit of data of the internal module 24-bit fixed-point multiplier module

was carried out above, here the analysis of power consumption data is done for the source code fmpy, the algorithm of adding booth and Wallace normally, and the final algorithm of the optimized boothwallace module to get the final power consumption statement. The different temperature process angles from 25c, 85c to 125c are shown in Tables 5–7. In addition to the data on the above aspects, this paper adds a record on the peak power, which helps to clarify the choice of chip power. In the integrated netlist section of this paper, the timing.sdc file is designed with an excitation of 0.5 ps, reaching the final measurement frequency of 2 GHz.

The combined and noncombined area comparison of the floating-point multiplier for the three algorithms is obtained in Figure 13, the total power consumption comparison in Figure 14, and the peak power comparison in Figure 15.

From the above tables, we can see that after the modification of the algorithm in the noncombination of the area is greatly reduced, although the combination of logic part compared to the most just fmpy algorithm way to increase, but adding combinatorial logic modules to the fixed-point

Table 5: 25c single-precision floating-point multiplier power consumption table.

2 GHz (TC-25C)	Area (nonportfolio)	Area (combination)	Total power	Peak power
fmpy	568.5504	375.9168	$3.016E-04$	0.1556
Naomal	71.1552	560.0448	$1.630E-04$	0.1493
boothwallace	67.1443	290.8974	$1.246E-04$	0.1432

Table 6: 85c single-precision floating-point multiplier power consumption table.

2 GHz (TC-25C)	Area (nonportfolio)	Area (combination)	Total power	Peak power
fmpy	568.5504	375.9168	$3.780E-04$	0.1586
Naomal	71.1552	560.0448	$2.133E-04$	0.1496
boothwallace	67.1443	290.8974	$1.562E-04$	0.1439

Table 7: 125c single-precision floating-point multiplier power consumption table.

2 GHz (TC-25C)	Area (nonportfolio)	Area (combination)	Total power	Peak power
fmpy	568.5504	375.9168	$3.192E-04$	0.1136
Naomal	71.1552	560.0448	$1.894E-04$	0.1025
boothwallace	67.1443	290.8974	$1.396E-04$	0.1012
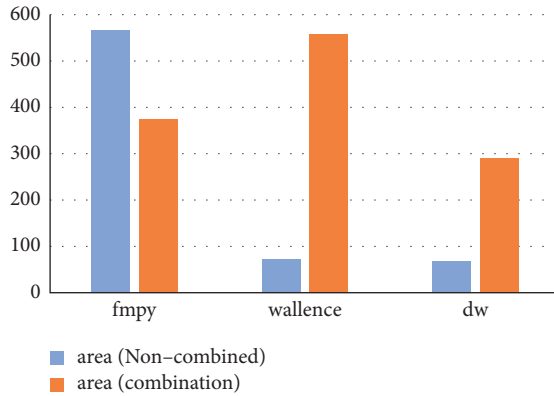
Figure 13: Area of combined and noncombined floating-point multiplier modules.
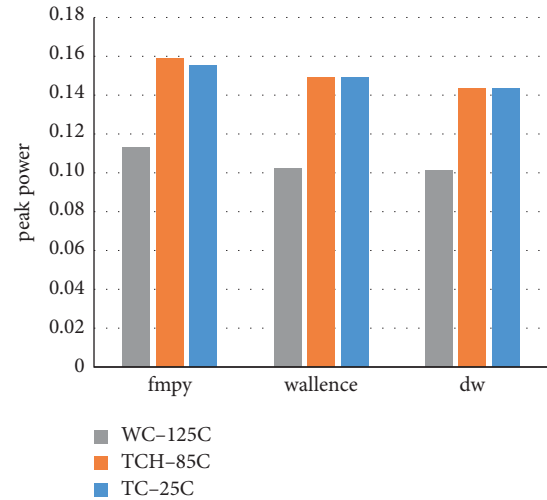
Figure 15: Comparison of floating-point multiplier spike power consumption at three temperature process angles.
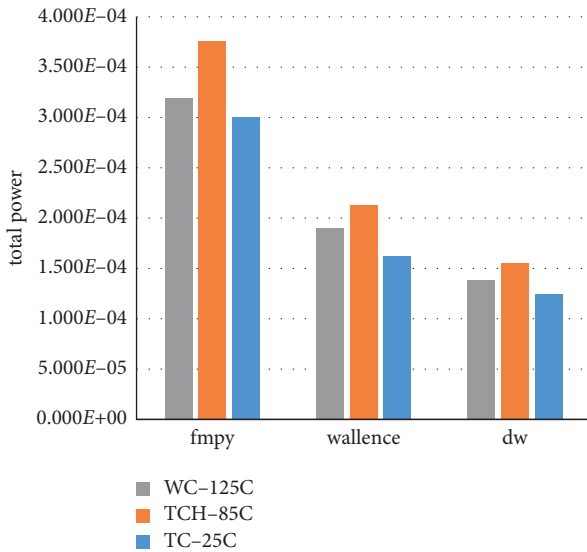
Figure 14: Comparison of total power consumption of floating-point multipliers at three temperature process angles.

modules will definitely create excess losses. The area of the floating-point algorithm is still much smaller than that of the fmpy algorithm of the shift-hosting method and the normal algorithm of simply adding the booth and Wallace modules. The total power of the modular algorithm of simply adding booth and Wallace is much lower than the original shift-hosting fmpy algorithm in the power analysis, with a reduction of 45.95%, while the optimized boothwallace algorithm module total power algorithm compared to the shift-hosting fmpy algorithm reduction of 58.69%, proving that optimizing Wallace will get very significant power savings, and on top of that adding a symbolic bit expansion module there is a not insignificant secondary improvement. In the peak power measurement, in various temperature process angles, the peak power reduction of the normal algorithm with the simple addition of the booth and Wallace modules and the optimized boothwallace algorithm compared to the
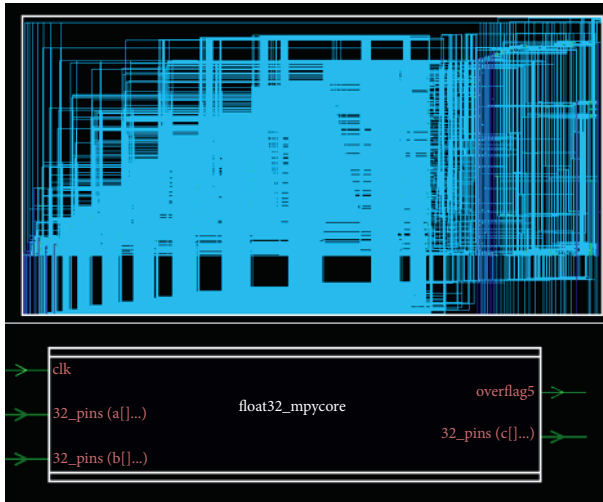
FIGURE 16: Integrated results' netlist diagram.

original shifted hosting fmpy algorithm are 3.79% and 7.97%, respectively, showing that the optimized algorithm also has a very significant improvement in peak power, which can make it a more generous choice in terms of power supply. At this point, the design fully meets the usage requirements.

The final result comes down to the resulting integrated netlist picture and cell diagram in Figure 16.

4. Conclusion

At this point in the paper, the single-precision floating-point multiplication algorithm has been logically designed and ultimately power-optimized through a series of rigorous verification and simulation processes. From source code writing, modification, verification on various platforms, synthesis, simulation, to the final experimental data, a comparative analysis of the experimental and control groups on multiple databases is performed. By using the symbolic bit expansion method to predict the partial product caused by the booth-4 code by rounding, the area of the resulting partial product tree is reduced. Then, the corresponding add operation and storage area are reduced in the covering structure of the Wallace tree, achieving the requirement of reducing the area and reducing power consumption. At this point, the algorithm design meets the expected requirements.

Data Availability

All data included in this study are available upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this study.

References

[1] C. Saran, *Stanford University Finds that AI is Outpacing Moore's Law*, Computerweekly. com, London, UK, 2019.

[2] A. Sodani and C. Processor, "Race to Exascale: Opportunities and challenges," in *Proceedings of the Keynote at the Annual IEEE/ACM 44th Annual International Symposium on Microarchitecture*, Porto Alegre, Brazil, December 2011.

[3] S. Yin, P. Ouyang, and J. Yang, "An energy-efficient reconfigurable processor for binary-and ternary-weight neural networks with flexible data bit width," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 4, pp. 1120–1136, 2018.

[4] L. Su, "Delivering the future of high-performance computing," in *Proceedings of the 2019 IEEE Hot Chips 31 Symposium (HCS)*, pp. 1–43, IEEE Computer Society, Hyderabad, India, December 2019.

[5] G. Marcus, P. Hinojosa, A. Avila, and J. Nalozco-Flores, "A fully synthesizable single-precision, floating-point adder/substractor and multiplier in VHDL for general and educational use," in *Proceedings of the IEEE International Caracas Conference on Devices*, IEEE, Punta Cana, Dominican Republic, November 2004.

[6] J. Ding and S. Li, "Determine the carry bit of carry-sum generated by unsigned MBE multiplier without final addition," in *Proceedings of the 2017 27th International Conference on Field Programmable Logic and Applications (FPL)*, pp. 1–4, Ghent, Belgium, September 2017.

[7] G. Haridas and D. S. George, "Area efficient low power modified booth multiplier for FIR filter," *Procedia Technology*, vol. 24, pp. 1163–1169, 2016.

[8] R. R. Desella, "Design and implementation of advanced modified booth encoding multiplier," *International Journal of Engineering Science*, vol. 2, no. 8, pp. 60–68, 2013.

[9] N. Bai, L. Wang, Y. Xu, and Y. Wang, "Design of a digital baseband processor for UHF tags," *Electronics*, vol. 10, no. 17, p. 2060, 2021.

[10] M. George, *Improved 24-bit Binary Multiplier Architecture for Use in Single Precision Floating Point Multiplication*, 2018.

[11] B. Rashidi, S. M. Sayedi, and R. R. Farashahi, "Design of a low-power and low-cost booth-shift/add multiplexer-based multiplier," in *Proceedings of the 22nd Conference on Electrical Engineering*, IEEE, Tehran, Iran, May 2014.

[12] C. G. Hooks, *Symbolic Expansion of Complex Determinants*, 2004.

[13] K. Bhardwaj and Mane, "Power- and area-efficient approximate wallace tree multiplier for error-resilient systems," *International System Quality Electronic*, 2014.

Hindawi

Research Article

Research on Athlete Behavior Recognition Technology in Sports Teaching Video Based on Deep Neural Network

XianPin Zhao [iD]

Wenzhou Polytechnic, Wenzhou Zhejiang, 325003, China

Correspondence should be addressed to XianPin Zhao; 2017091818@wzpt.edu.cn

In recent years, due to the simple design idea and good recognition effect, deep learning method has attracted more and more researchers' attention in computer vision tasks. Aiming at the problem of athlete behavior recognition in mass sports teaching video, this paper takes depth video as the research object and cuts the frame sequence as the input of depth neural network model, inspired by the successful application of depth neural network based on two-dimensional convolution in image detection and recognition. A depth neural network based on three-dimensional convolution is constructed to automatically learn the temporal and spatial characteristics of athletes' behavior. The training results on UTKinect-Action3D and MSR-Action3D public datasets show that the algorithm can correctly detect athletes' behaviors and actions and show stronger recognition ability to the algorithm compared with the images without clipping frames, which effectively improves the recognition effect of physical education teaching videos.

1. Introduction

With the development of society, more and more sports teaching videos have entered people's daily life. The analysis of PE teaching video can more effectively improve the teaching effect. Object detection and human pose estimation for sports video are the basis of sports video analysis and understanding. The existing target detection and human pose estimation technologies have achieved good performance in the general scene detection task based on pictures [1], but there are few algorithms and data for target detection in sports video scenes. For a new data field, accurate athlete detection and pose estimation are important links in sports video analysis. The existing human target detection and pose estimation algorithms have achieved good performance in the general human body detection task, but they will detect athletes and spectators at the same time in the physical education teaching video, so they cannot further distinguish athletes' targets, which will interfere with the subsequent video analysis. At the same time, the sports video data with human body annotation is scarce [2], and the cost of obtaining a model suitable for the field of sports video is

high. Athletes are a special case of general human body detection task. If the general human body detection model can be used to detect and estimate athletes in sports video, it can undoubtedly save a lot of cost. In order to reduce the labelling cost and training cost, the detection and pose estimation of athletes in sports teaching video are realized.

Among the traditional methods, the two-dimensional behavior recognition technology based on RGB video has been widely studied [3]. The early method is to extract local spatiotemporal descriptors from the input video and encode these descriptors into word bags with visual vocabulary for classification. At the same time, the local feature method has become an effective way in motion recognition, because these local features do not need algorithms to detect the human body and are not sensitive to factors such as background, illumination, and noise. The calculation of these local features can usually be divided into two parts. The calculation of these local features can usually be divided into two parts: detector and descriptor. Detector is to find the significant and informative region of action understanding and descriptor, and descriptor is to describe the visual model of the extracted region. Laptev proposed spatiotemporal

points of interest by extending Harris angle to spatiotemporal dimension [4], and Kviatkovsky et al. proposed a covariance descriptor to realize online action recognition [5]. Local descriptors for video representation include histogram of oriented gradient (HOG) features and histograms of optical flow (HOF) features [6]. The two-dimensional convolution neural network used for RGB static pictures cannot meet the requirements of extracting time features. Another method is to apply the three-dimensional convolution neural network, but the three-dimensional neural network has a great disadvantage that there are too many parameters to be optimized, so it needs a large-scale dataset for training [7].

Inspired by the successful application of depth neural network based on two-dimensional convolution in image detection and recognition, this paper takes physical education video as the research object and constructs a depth neural network based on three-dimensional convolution to automatically learn the temporal and spatial characteristics of human behavior, which is used for human behavior recognition. Combining convolution neural networks and recurrent neural networks, convolution neural networks are used to extract spatial features, and then cyclic neural networks are used for time series modeling. This method can avoid the demand for optical flow and better meet the real-time requirements. Taking the deep learning method as the key technology, VGG16 convolutional neural network is used to extract the highly abstract features of video frames, and an athlete behavior recognition algorithm in sports teaching video is designed and implemented.

2. Related Work

At the beginning of this century, some foreign universities began to pay more attention to the research of human behavior recognition. Today, a lot of research has been carried out in the field of human behavior recognition [8], and more and more valuable research results have emerged. The recognition methods have made rapid progress and changed with each passing day. With the rise of deep learning, the research on human behavior recognition using deep neural network has begun to be more and more in depth. Compared with foreign countries, China started late in this area, and there is still a certain gap to catch up. In recent years, it has begun to strengthen research. Davis and Taylor [9] first used contour to describe the global feature of human body, which is a two-dimensional feature. Motion energy image (MEI) is used to indicate the position where motion has occurred, motion history image (MHI) is used to indicate the spatial position and time sequence of motion, and then recognized by Mahalanobis distance classifier, these two kinds of feature images are obtained by background frame difference method. The optical flow method based on pixel level can capture and locate the moving object. When the background frame difference method cannot extract the foreground object well, the optical flow method can be used for recognition and location. It is easy for the optical flow method to introduce the motion noise of noninterested objects. Literature [10] only calculates the

optical flow at specific positions (such as the approximate center point of the object), which reduces the influence of noise to a certain extent.

Blank et al. [11] first obtained 3D spatiotemporal volume (STV) from silhouette information in video or picture sequence, then derived local spatiotemporal interest points and their directional features through Poisson equation, and obtained a 3D global feature by weighting these local points. In contrast to the global feature, the local feature adopts the bottom-up research thinking, and only relatively independent image blocks in the object region are extracted. Since the region of interest (ROI) [12] does not have to be a complete object with practical significance in the real world, accurate object region positioning and tracking are not required. The usual approach is to calculate some spatiotemporal interest points in the video or picture sequence, then expand and fuse them into image blocks one by one, and finally combine all image blocks to generate total features. Because it does not depend on the accurate positioning and tracking of the bottom layer, the sensitivity of local features to noise, angle change, and occlusion is not high. However, due to the need for a large number of stable interest points related to action categories, various complex preprocessing processes cannot be avoided. At the same time, the motion of the camera will also affect the calculation and detection of interest points. In the traditional human behavior recognition algorithm, it is necessary to artificially control the complex process of feature extraction and data reconstruction. By using the deep neural network, the picture can be directly used as the input of the network, and the recognition effect can be better than that of the traditional method [13]. Deep neural network does not completely break away from the traditional methods and start a new stove but integrates the complex steps of the traditional methods into the network and simplifies the whole process. All connections and parameters in the network reflect some ideas of the traditional methods. Compared with the traditional shallow neural network, deep neural network, as its name implies, has more network layers in structure and a large number of parameters, and its internal connection mode will change greatly with the different types of networks. The multilayer structure of the network enables it to learn higher-level and more abstract expressions [14], so that the features extracted from the input data will have higher discrimination.

Human posture estimation is also called human joint point regression, which detects the position of each joint point of the human body and then connects it to the human skeleton. Human posture estimation is not only the basis of behavior analysis and motion recognition but also an important part of video analysis [15]. If the human joint points are returned to 3D for expansion, RGB images are input, and 3D human key points are output, it is 3D human pose estimation. This is a very valuable research direction. The traditional human posture estimation algorithm divides the human body into multiple parts according to joints, hand, head, leg, and body and then performs feature matching based on the manually extracted features of each part to detect the limbs in the figure. Taking the body as the trunk,

the detected limb parts are spliced according to the Gaussian model of human posture distribution to obtain a complete human posture [16]. This method needs to manually design human features, and the feature engineering steps are complex. When the scene is complex, it is difficult to cover all the situations, which has great limitations and cannot be effectively popularized. With the development of deep learning, human pose estimation algorithm based on deep learning has become the mainstream. The human posture estimation algorithm based on deep learning regards human posture detection as a key point regression problem, trains through a large number of data with joint point category and position annotation, and finally obtains a model that can predict the position and category of human joint points [17].

The existing deep learning models mainly include convolutional neural networks (CNNs), recurrent neural networks (RNNs), restricted Boltzmann machine (RBM), and deep belief network (DBN) [18]. These methods give full play to the self-learning ability of neural network and let the network adapt to the characteristics of data. The method of manually selecting features is avoided. A large number of studies show that these models can achieve better results than the original methods in handwritten numeral recognition, target detection and recognition, speech recognition, and other scenes.

2.1. Athlete Behavior Recognition Model Based on Deep Neural Networks

2.1.1. Athlete Behavior Tracking. There are strong rules in shooting and editing techniques in sports video. The characteristics of different shots and the differences between different shots are obvious. In the far shot, the competition field accounts for the largest proportion, and the athletes are also concentrated on the competition field. Therefore, the competition field area is detected first. After the athletes are detected, the more accurate athletes can be obtained by removing the field color by using the consistent color characteristics of the sports field. In the middle lens, the field color still accounts for a certain proportion in the image, so if the far lens scene has been processed, the main field color is stored to remove the nonathlete area of the image belonging to the competition field.

Athlete detection is a classifier that combines the characteristics of athletes and background in various types of shots in sports video and uses middle-level feature blocks to train and distinguish athletes. The middle-level characteristics of neural network often contain rich information, which plays a connecting role in the whole neural network. In this paper, CamShift algorithm is used to track athletes, and its core algorithm is mean shift algorithm. Mean shift algorithm is a nonparametric probability density estimation method. Its algorithm block diagram is shown in Figure 1. It finds the peak points in the probability distribution through the iterative process and makes each point "drift" to the local maximum of the density function. CamShift algorithm is an extension of mean shift algorithm. It realizes the tracking of moving objects in continuous video frames by adaptively adjusting the size and position of search window. It has the characteristics of good robustness, fast speed, and less computation. Using kernel function estimation method, when the sampling is sufficient, it can gradually converge to any density function; that is, it can estimate the density of data subject to any distribution. In this way, the convergence of mean shift algorithm can be guaranteed, and the mean shift vector can be used for iteration. It can better track the irregular motion of nonrigid objects and has good antinoise ability.

An important step in CamShift algorithm is to convert the original video image into a probability distribution map by using the backprojection of the color histogram of the target object. Each pixel value in the map represents the probability that the pixel belongs to the target object. Because HSV color space separates color (H) from saturation (S) and brightness (V), using the H channel of HSV color space to establish histogram can reduce the sensitivity to illumination changes. The traditional method of calculating histogram is as follows:

$$q_u = \sum_{x \in R} \sum_{y \in R} c\delta(x, y) - u^2, \quad (1)$$

where $\{q_u\} u = 1, 2, \ldots, m$ is the histogram; R is the target area where histogram needs to be calculated; the function C: $R_2 \longrightarrow \{1, 2, \ldots, m\}$ is used to convert the pixel value with coordinates of (x, y) into the histogram level value; and δ is the Kronecker symbol:

$$\delta[i - j] = \begin{cases} 0, \text{other} \\ 1, i = j \end{cases} \quad (2)$$

In order to make the obtained probability distribution map within the range of [0255], the histogram needs to be scaled:

$$p_u = 255 \cdot \min\left(q_u q_{\max}^{-1}, 1\right), \quad (3)$$

where $u = 1, 2, \ldots, m$; $q_{\max} = \{\max(q_u)\} u = 1, 2, \ldots, m$. Let the value of the pixel at (x, y) in the video image be $I(x, y)$ and p be the probability distribution map corresponding to the image, then, $P(x, y) = P_{f[I(x,y)]}$, where $\{P_u\}, u = 1, 2, \ldots, m$ is the histogram; function $f[I(x, y)]$ finds the corresponding level value in the histogram according to the value of $l(x, y)$.

CamShift is a semiautomatic tracking algorithm, which needs to explicitly specify the initial size and position of the search window. In order to realize automatic tracking, this paper takes the detected athlete area as the initial search window. In the iterative process, the i centroid (x_c, y_c) of window w is

$$x_c == \sum_{x \in w} \sum_{y \in w} x P(x, y) \left[\sum_{x \in w} \sum_{y \in w} P(x, y)\right]^{-1},$$

$$y_c = \sum_{x \in w} \sum_{y \in w} y P(x, y) \left[\sum_{x \in w} \sum_{y \in w} P(x, y)\right]^{-1}. \quad (4)$$

Each iteration resizes the window to $2^{-3} \left[\sum_{x \in w} \sum_{y \in w} P(x, y)\right]^{1/2}$.
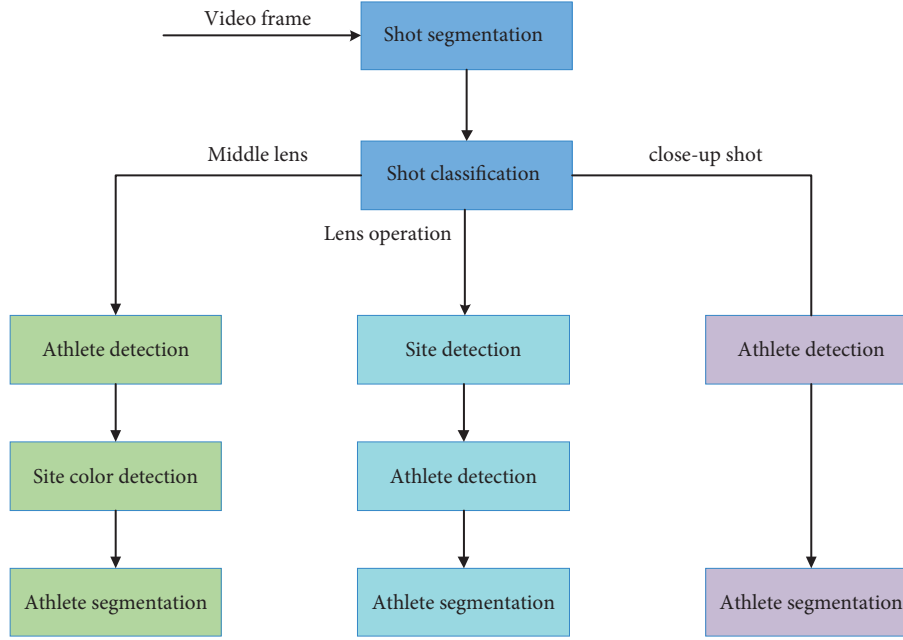
Figure 1: Monitoring and segmentation of athletes in sports teaching video.

Because the traditional histogram only uses the information of H channel, it is easy to fail to track when the hue distribution of the target object is similar to that of the background.

In order to make CamShift's tracking algorithm more robust, this paper uses the weighted histogram method to improve the algorithm. In the athlete's area, the pixel farther away from the center of the area has lower reliability, and the pixel is more likely to be covered by other objects or belong to the background. Therefore, it is reasonable to assign a relatively low weight to the pixel far away from the center of the area when calculating the histogram. The weighted histogram is calculated as follows:

$$q_u = \sum_{x \in w} \sum_{y \in w} k(r) \left[c\delta(x, y) - u^2 \right], \qquad (5)$$

where $K(r)$ is a monotonically decreasing function, (x_c, y_c) is the regional center, and $r = \sqrt{(x - x_c)^2 + (y - y_c)^2}$. In order to ensure the real-time tracking, this paper sets the number of iterations of each mean shift as 15 times, and the convergence standard is to drift one pixel.

2.1.2. Behavior Feature Extraction Based on VGG16 Model.
The feature extraction layer of VGG16 convolution neural network model is used to process the frame picture. The input of the model is the cropped frame picture, which needs to be normalized to 224×224; this paper takes the output of the feature extraction part, which is a 4096-dimensional vector. Each picture in the frame sequence needs feature extraction and finally forms a feature sequence.

VGG networks are divided into several classes according to the number of layers. The structures of all classes are very consistent, and almost all convolution layers use 3×3 size convolution kernel (only one class uses 1× the convolution

kernel with the size of 1 being regarded as a linear mapping of space), and the pooling layer also adopts 2×2 size aggregation. Using a smaller convolution kernel can reduce a considerable part of the parameters under the action of weight sharing. In shallow convolution neural network, less parameters often lead to insufficient generalization ability of the model, resulting in underfitting. At this time, the network parameters can be effectively increased and the expression ability of the model can be improved by superimposing convolution layer. In addition, it also creates conditions for more nonlinear mapping and enhances the fitting ability of the network. Common VGG networks usually have 16–19 convolution layers. There are obvious groups between these convolution layers. At the end of each group, there is a pool layer, and the output of this group is downsampled. VGG16 network model is adopted in this paper, with a total of 16 convolution layers. Its structure is shown in Figure 2.

In Figure 2, there are 5 groups of convolution layers, and the size of convolution kernel is 3×3. The first group has two convolution layers, each layer has 64 convolution cores to extract features from multiple angles and dimensions, the second group has two layers, 128 convolution cores per layer, the third group has three layers, 256 convolution cores per layer, and the fourth and fifth groups also have three layers, 512 convolution cores per layer. Each group of convolutional layers is followed by a 2×2 pooling layer. At the end of the network is a 3-layer fully connected layer. The entire network has a total of 138 million parameters to be trained.

In terms of training, VGG16 will first perform necessary preinitialization of some layers, including include_top: whether to include the top fully connected layer; weights: None stands for random initialization of "ImageNet," which stands for pretrained weights loaded on ImageNet;
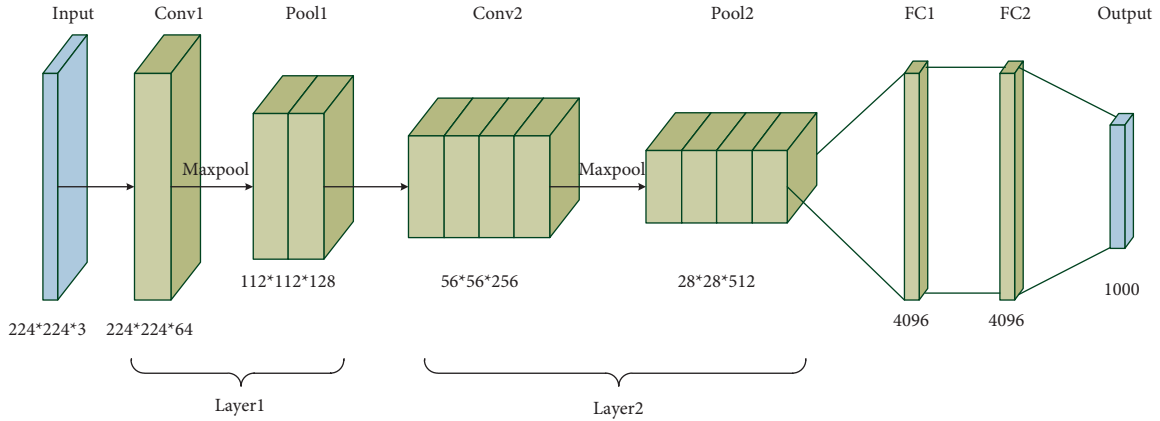
FIGURE 2: VGG16 network model structure diagram.

input_tensor: Optional, Keras tensor is used as the input of the model (that is, the tensor output by layers.Input()); input_shape: Optional, input size tuple, valid value when only include_top = False; pooling: Optional. When include_top is False, this parameter specifies the pooling method for feature extraction. In addition, the depth of the network and the small size of the convolution kernel have the effect of implicit regularization, so it only needs a few iterations to start convergence. This paper uses the VGG16 model fully trained on ImageNet, because its input scale is fixed at 224×224, so the pictures in the frame sequence need to be reduced to the corresponding size before they can be sent to the network. The length width ratio of the cropped frame picture is about 219×214, so there will not be too much distortion after zooming. In addition, the network itself has certain scale invariance, so it can be put into the network for feature extraction.

In terms of calculation, as mentioned above, the convolution layer mainly convolutes or performs correlation operations on the pictures, the pooling layer mainly subsamples each feature image obtained from the previous layer to reduce its size and retain useful information while reducing the amount of data, and the full connection layer integrates local features to achieve higher abstraction or data classification or prediction (the last layer is the full connection layer). Like most deep neural networks, the activation units of all convolution layers in the network adopt the modified linear unit (ReLU) activation function. The activation function of artificial neuron generally adopts nonlinear monotone function which is continuous and differentiable everywhere, and the value range of the function is a fixed interval. The requirement of nonlinearity is that the linear function can only play the role of linear amplification of the input, and the linearity will greatly reduce the performance of the network, and it is easy to degrade the multilevel network function into a single-level network function, which is a fatal defect. The training algorithm can calculate the gradient and adjust the parameters by using the gradient. Monotonic performance can effectively improve the efficiency of network training. The bounded value range that can be given by the actual demand can make the gradient decline more stable and prevent the

network from entering the saturation state: if the input is generally small, the value range can be adjusted to have a larger gain. If the input is generally too large, adjust the value range to get a smaller gain.

The activation effect of ReLU function is similar to that of biological neurons, and its mathematical expression is shown as follows:

$$f(x) = \begin{cases} 0 & x < 0, \\ x & x \geq 0. \end{cases} \tag{6}$$

As shown in Figure 3, the blue line and red line, respectively, represent the convergence effect of ReLU activation function and sigmoid activation function. It can be seen from the figure that the convergence speed of training model with ReLU activation function is significantly faster than that with sigmoid activation function, which greatly shortens the learning cycle.

Deep learning is actually a feature driven method. As a typical method of deep learning, the network model composed of convolutional neural network is no exception. Usually, the output of the feature extraction layer of the network can be taken out separately as a feature description of the data, and these feature descriptions can be used to classify or predict the data in other ways. In VGG16, after layer by layer calculation, the input image gradually transitions from local feature extraction to global feature construction. The output of each layer can be taken out separately as a feature description. The last three layers of the network are fully connected to the network layer, except the last layer (used to directly process the features extracted by the previous layer to classify or predict the data) that can well describe the global features of the image. Take the output of the first fully connected layer, whose output form is 4096-dimensional vector, and the extracted feature description will be used as the input of the next layer network.

2.1.3. Training Method of Athlete Classifier in Sports Teaching Video. In sports video, the characteristic of athletes and background is that all pictures contain background and athletes, and there are basically no frame images without athletes. In addition, in addition to close-up shots, athletes
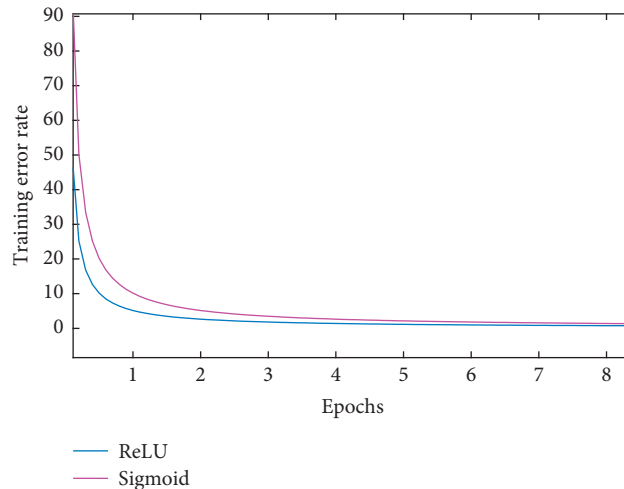
FIGURE 3: Comparison of convergence effect.

account for a small proportion of the frame images in other shots, and most of the frame images are background blocks. Moreover, the background in the sports scene is the competition venue, and the competition venue structure of the same type of sports competition is very similar. In this case, it is not feasible to mine middle-level feature blocks by unsupervised iterative clustering detection. In the initial clustering, due to the large number of background blocks and unified structure, most of the classes with less than or equal to three image blocks in the discarded class are foreground blocks. The cohesive clustering algorithm makes decisions according to the local pattern without considering the global structure of the data. Therefore, cohesive clustering is very suitable for identifying smaller classes. When using unsupervised iterative clustering detection method, the first thing to do is to separate the background from the athlete block. Because the sample collection stage can only be separated manually, the workload is too large.

From another point of view, these characteristics of athletes in sports video can simplify the definition of middle-level feature blocks. In the close-up shot, the athlete only has a bust, and there are not many types of athletes. There are not many kinds of blocks that contain different parts of athletes. In the middle shot and far shot, the proportion of athletes in the image is small; especially in the far shot of football, tennis, and badminton, most athletes can use a small block for complete segmentation. The number of blocks with large difference between athletes in the image segmentation of middle lens and far lens is not very much. Based on this feature, the middle-level feature block that can distinguish athletes and nonathletes is defined as the block with the most athlete contour. All qualified blocks are used as the positive samples of the training athlete classifier, and other blocks are used as the negative samples of the training athlete classifier. The following will explain how to collect positive samples and negative samples to train the athlete classifier in the sports video.

The steps of collecting positive and negative samples of the classifier for training athletes are shown in Figure 4. It should be noted that, in order to calculate the most similar

blocks and the K-means clustering results in the classification results more accurately, the pictures of each segmented block are from the same video. The image is segmented with a fixed window and step size to obtain the image block set. The image blocks with the largest athlete contour and different parts are selected as seed blocks, and the k most similar blocks and k least similar blocks of each seed block are calculated, respectively. Taking all the blocks most similar to the seed block as the positive sample set and the least similar block as the negative sample set, the hog features of all positive and negative sample blocks are extracted, and the SVM classifier is used to train the athlete's classifier, which is the initial athlete classifier. In the later iteration, the classifier is used to classify and detect all blocks, and then K-means is used to cluster the pictures with positive and negative results. Manually check whether the classification corresponding to each clustering result is wrong. If it is wrong, add the wrong class to the correct sample, so as to expand the positive and negative sample set. At the beginning of the next iteration, reextract the features of the sample set and train a new classifier, which optimizes the classifier. Iteratively detect, cluster, and update the optimized classifier until the classifier correctly classifies a certain number of new video picture sets.

2.1.4. Athlete Behavior Recognition Model and Overall Recognition Process Based on Deep Neural Network. Combining the feature extraction of spatial dimension and the information integration of time dimension, we combine to form a VGG16-BiLSTM athlete behavior recognition model, as shown in Figure 5.

The model takes the cropped frame sequence as the input and the output vector of the last time of the cyclic neural network as the result. The first part of the model is the feature extraction layer of VGG16 deep convolution neural network, which adopts the weight parameters fully trained on ImageNet and takes the output of its first full connection layer as the extracted image features. The second part is the bidirectional long-term and short-term memory model with
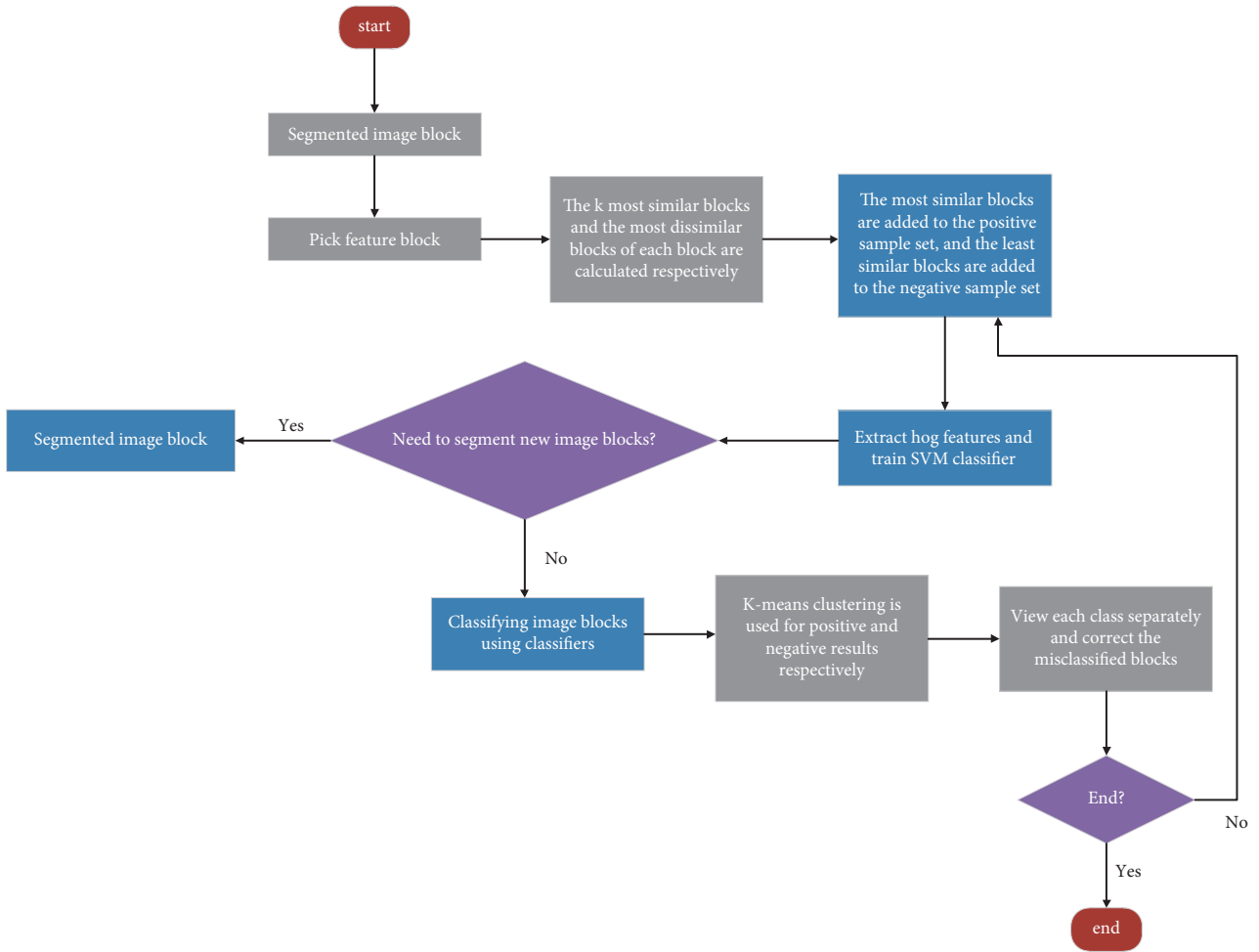
FIGURE 4: Process of training athlete classifier.
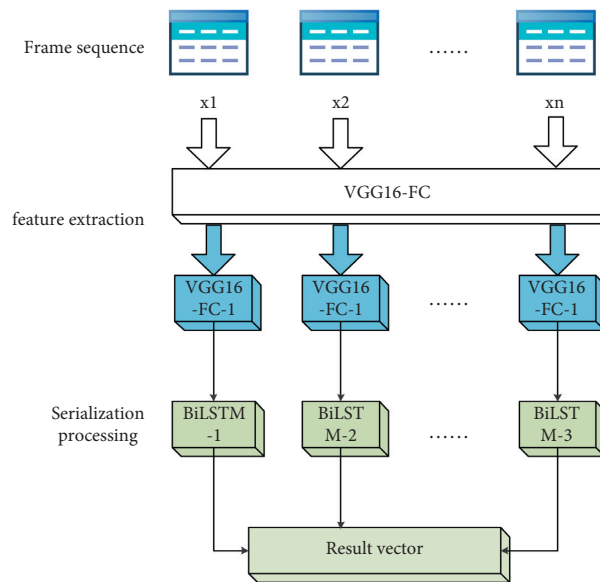
FIGURE 5: Model diagram of deadlock detection method for hierarchical scheduling system based on time constraints.

more advantages in integrating time dimension information. Its weight parameters are trained with the feature sequence extracted from the real shooting data of physical education teaching video used by the subject as the training data. It takes the feature sequence obtained in the first part as the input and outputs the recognition results of the whole model. Integrating all the above research contents, Figure 6 shows the overall identification process in the final athlete behavior detection system.

3. Experiment and Analysis

3.1. Experimental Setup and Pretreatment. Two public datasets, UTKinect-Action3D and MSR-Action3D, are used to evaluate the proposed method. They are the behavior video data of physical education teaching video captured by Kinect. There are 10 behaviors in UTKinect-Action3D, namely, walk, sit down, stand up, pick up, carry, throw, push, pull, wave hands, and clap hands. There are 10 PE teaching videos, each behavior in each teaching video twice, and a total of 199 effective videos. For the convenience of calculation, all 200 videos are used. There are 20 behaviors in MSR-Action3D dataset, which are shot by 10 subjects, in which each athlete completes each behavior 2-3 times.

According to the experimental settings in the basic project research of the PE teaching video dataset, 20 behaviors are divided into three behavior subsets, AS1, AS2, and AS3 (as shown in Table 1). Each behavior subset contains 8 different behaviors.

In order to reduce the impact on different experimental results, each video is simply preprocessed before the experiment:

(1) Background removal: in depth video, the depth information of the background is consistent, while the depth information of the foreground is changed. The background information can be removed according to this feature. Gaussian model is a common method of background removal in video image detection. This method is suitable for dynamic video image feature detection, because the foreground and background are separated in the model. The benchmark for separating foreground and background is to judge the change rate of pixels. The learning with slow change will be regarded as the background, and the learning with fast change will be regarded as the foreground.

(2) Determination of bounding box: for each video, the bounding box that can and can only frame the athlete's behavior is obtained according to each frame, and the maximum bounding box of all frames is taken as the bounding box of this video (see Figure 2).

(3) Normalization: use interpolation technology to normalize all videos processed in the previous step to a unified size, where the number of normalized video frames is equal to the middle value of all video frames. At the same time, the min max method is used to normalize the depth information values of all

videos to the range of [0, 1]. Finally, all samples are turned horizontally to form new samples, so as to multiply the training samples in the dataset. After preprocessing, the behavior video sizes of UTKinect-Action3D and MSR-Action3D are $28 * 32 * 32$ and $38 * 32 * 32$, respectively, including the number of frames, frame width, and frame height in the video from front to back. The experiment depth neural network model part is compiled by torch platform, and the data preprocessing part is completed by Matlab platform.

3.2. Performance of Athlete Behavior Recognition Algorithm Based on Deep Neural Network. Firstly, the effectiveness of the research method is verified on the MSR-Action3D dataset. According to the experimental setting of [19], the research method is compared with the benchmark project research of the dataset [19] and several main methods based on artificial feature extraction in recent years. Figure 7 shows the method and the accuracy of behavior recognition on three different behavior subsets. From the recognition results, it can be seen that the athlete behavior recognition method based on 3D convolution depth neural network can effectively recognize human behavior, and the recognition accuracy and average accuracy of each behavior subset are better than the benchmark research of the dataset.

This is mainly because the 3D word bag model is used to extract the features in the behavior video. This feature can extract the representative 3D word bag information in the physical education teaching video but ignores the spatial and temporal information in the video. When doing video evaluation experiments, we need to select representative video sequences according to certain conditions. Temporal information (Ti, temporal perceptual information, also known as temporal complexity) and spatial information (Si, spatial perceptual information, also known as spatial complexity) can be used to measure the characteristics of video. Where Si is high, there are more spatial details of video frames (high spatial complexity). Where Ti is high, the video frame content moves violently (with high time complexity). The athlete behavior recognition method based on 3D convolution depth neural network adopts 3D convolution operation for the video, which effectively maintains the spatial and temporal features. Therefore, better performance is obtained. Table 2 shows the comparison between the research method and the current method with the best recognition efficiency [20]. The experiment adopts the setting of the method in [20]. The results show that the proposed method can maintain consistent recognition performance with the methods in [20], which fully shows the effectiveness of this method (Figure 8).

In the UTKinect-Action3D dataset, the research method is compared with the benchmark research project on the dataset. Literature [21] used leave-one-out cross validation method (LOO-CV). For the convenience of the experiment, the study uses leave one subject out cross validation (LOSO-CV); that is, only one subject's all behavior videos are used as the test set at a time, while the data of other subjects are used
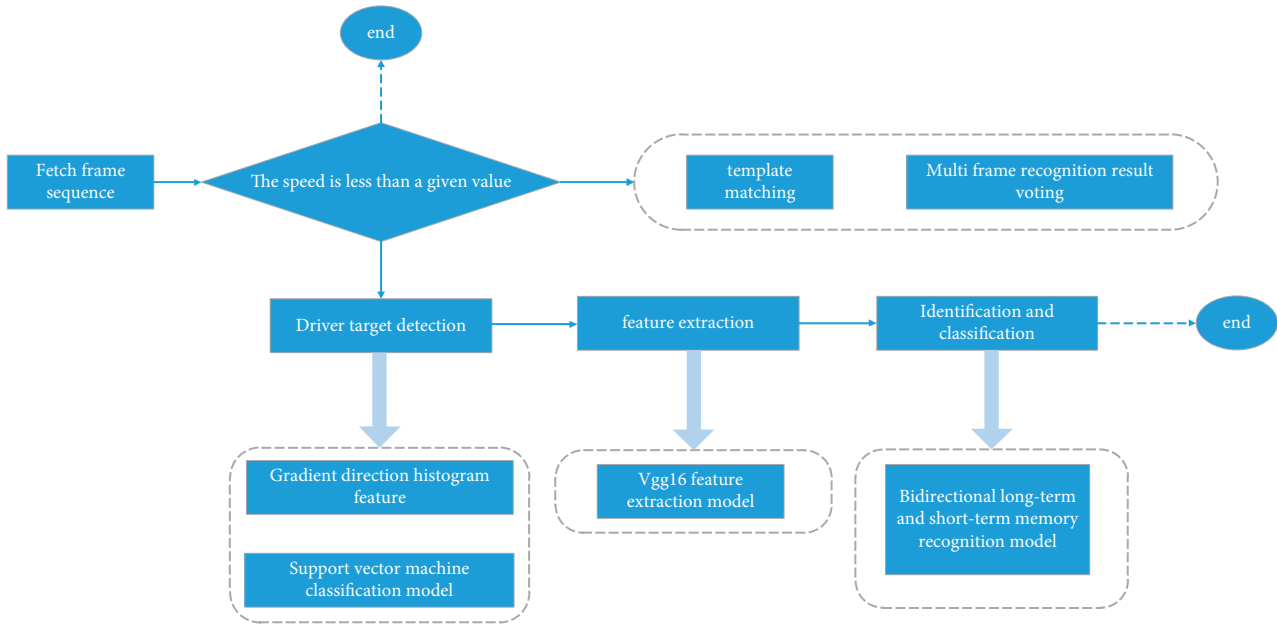
FIGURE 6: Flowchart of the identification process in the final athlete behavior detection system.

TABLE 1: The action subsets AS1, AS2, and AS3 in MSR-Action3D dataset.

AS1	AS2	AS3
Horizontal arm wave	High arm wave	High throw
Hammer	Hand catch	Forward kick
Forward punch	Draw x	Side kick
High throw	Draw tick	Jogging
Hand clap	Draw circle	Tennis swing
Bend	Two-hand wave	Tennis serve
Tennis serve	Forward kick	Golf swing
Pickup & throw	Side boxing	Pickup & throw
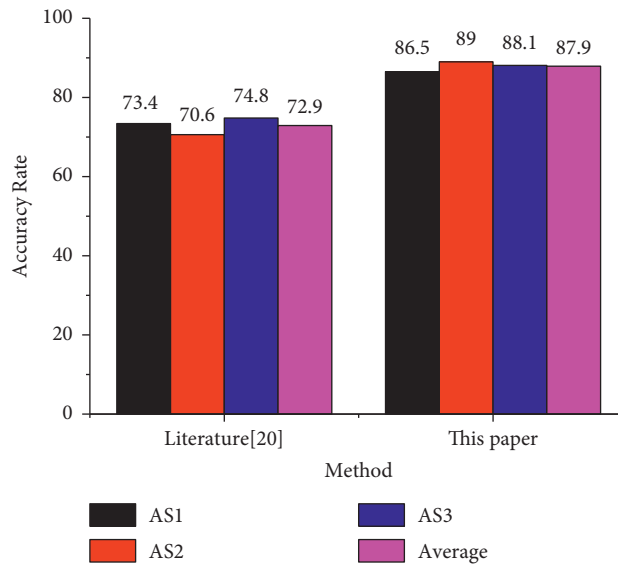
FIGURE 7: Comparison with the benchmark project in MSR-Action3D dataset (%).

TABLE 2: The action recognition accuracy of each subject in UTKinect-Action3D.

Subjects	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8
Recognition rate%	87	85	89	81	76	72	78	73

FIGURE 8: Performance evaluation compared with [20, 21] (%).

extraction links. It only needs simple processing of the original video, and the deep neural network model can automatically extract features and complete the recognition and classification process. The method is simple and involves fewer links.

3.2.1. Behavior Recognition Parameter Experiment. If the extracted frame image sequence is directly sent to the network model for training, recognition, and classification, the recognition effect will not be ideal because of too many interference factors. Direct training can not model complex video, and the video modeling effect with long time is not good. Although the recognition subject in the video is an athlete, other moving objects appear in the picture due to the angle and complex actual scene during video shooting, and their dynamic changes will interfere with behavior recognition. Therefore, in the process of preprocessing the data, the support vector machine is used to detect the gradient direction histogram characteristics of each region in the frame picture to locate the athlete's region and then cut off the picture information outside the region of interest. After the model is trained, the multiscale detection method is used to detect the target in a picture. The multiscale method uses the image pyramid and sliding window, which involves two important parameters: the downsampling scaling factor scale: the divisor of the width and height of the image divided by each downsampling.

Sliding step means that how many pixel units the sliding window slides each time it slides horizontally or vertically. The variation range of detection scale directly affects the detection effect and consumption time. Setting these two parameters reasonably is a key to optimize the recognition effect and recognition time. In the sequence data, the "sliding window" is used to intercept sequence fragments, thereby reshaping the original data into samples of a specified length for model modeling. On the premise of achieving the best detection effect, we should reduce the scale range as much as possible to reduce the detection time.

100 frame pictures from different videos are selected as the test data. Taking the ability to correctly detect the athlete's area and further obtain the athlete's area as the standard for successful recognition, two variables are controlled for detection, respectively. The detection results are shown in Figure 9, in which the ordinate represents the scale factor, and the abscissa represents the sliding step of the sliding window in the form of (horizontal sliding step, vertical sliding step).

It can be seen from the table that the average detection time is inversely correlated with the scaling factor and sliding step size. Considering the data in Figures 9 and 10, it is finally decided to select the scaling factor of 1.2 and the sliding step size of (16, 16).

as the training set, so as to train a deep neural network model for each subject. Obviously, the experimental conditions are more stringent than those in [20].

It can be seen from Table 2 that the average accuracy rate of behavior recognition of each subject is 80.125%, which can basically correctly identify the behavior of most subjects. However, the recognition accuracy rate on physical education teaching videos 5, 6, 7, and 10 is relatively low, which may be due to the large deviation of perspective of athletes in these sports teaching videos when shooting action behavior because UTKinect-Action3D is a multiperspective dataset. However, there is still a certain gap between the research method and the 90.92% recognition rate in [21], with a difference of nearly 8 percentage points. It may be because it makes full use of the skeleton information in the depth video and uses hidden Markov model (HMM) to establish the temporal model of skeleton information. Its disadvantage is that the behavior recognition framework is too complex, and the system performance is affected by skeleton information extraction and HOL3D feature extraction, feature LDA projection, behavior word clustering, HMM model training, and other links, and literature [21] can see that skeleton extraction is a complex process, and the accuracy of the extracted skeleton information depends on the shooting of depth video. At the same time, the experimental conditions of the research method are more stringent than those in [20], the experimental data are relatively small, and the training of the model is insufficient, which are also the reasons for the poor recognition effect. Nevertheless, compared with the methods of feature extraction such as [21], the method based on deep learning has better generalization performance, and the research method does not need complex artificial feature

FIGURE 9: Accuracy rate of multiscale detection (%).

FIGURE 10: Average time of multiscale detection (ms).

The experimental data are 250 training videos and 60 test videos. All videos take 30 frames at equal intervals and then use VGG16 network model to extract features and train and test through two-way long-term and short-term memory network model. Under the control of other variables, the frame sequences before and after clipping are used,

respectively, for experiments; the contrast effect before and after clipping is shown in Figure 11. After cutting, the overall recognition effect has been significantly improved, the recognition rate has increased from 42% to 27%, the single category false alarm rate has decreased, and the average false alarm rate has decreased by nearly half, as low as 9%.
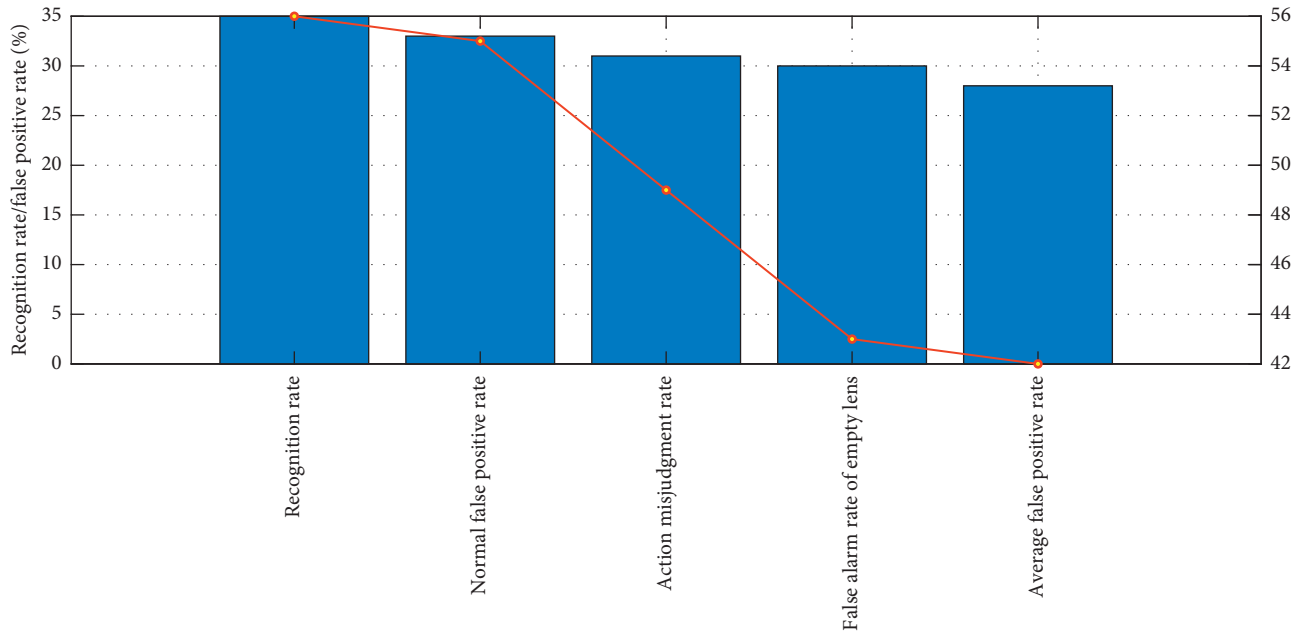
FIGURE 11: Comparison of recognition effect before and after cutting.

4. Conclusion

This research is based on a deep neural network and uses the CamShift algorithm to track athletes. It collects positive and negative samples of the training athlete's classifier and extracts the HOG features of all positive and negative sample blocks. CamShift can effectively solve the problem of target deformation and occlusion. Resource requirements are not high, time complexity is low, and good tracking results can be achieved in a simple background. However, when the background is more complicated, or there are many pixels that are similar to the target color, it will cause the tracking to fail. Therefore, the SVM classifier is used to train the athlete's classifier and uses the algorithm of cutting frame pictures and VGG16 model to recognize athletes' behavior. After some interference information in the picture is cut out in the data preprocessing of the system, in order to further strengthen the attention of the network model to the actions of athletes' key parts, the mechanism of adding attention in the cyclic neural network can be used to further improve the recognition effect. Although some progress has been made in improving model reusability and reducing labelling and training costs, there are still many aspects that can be improved, such as improving the speed of sports video target detection and human posture estimation. The adjacent frames in the video have very high similarity. This paper uses this to fuse the multiframe results, which improves the detection accuracy, but also brings a lot of redundant detection. In the later stage, we can consider accelerating the detection according to the redundant information between frames.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] A. Jalal, I. Akhtar, and K. Kim, "Human posture estimation and sustainable events classification via pseudo-2D stick model and K-ary tree hashing," *Sustainability*, vol. 12, no. 23, p. 9814, 2020.

[2] B. Fasel, J. Spörri, P. Schütz, S. Lorenzetti, and K. Aminian, "An inertial sensor-based method for estimating the athlete's relative joint center positions and center of mass kinematics in alpine ski racing," *Frontiers in Physiology*, vol. 8, p. 850, 2017.

[3] Y.-H. Byeon, D. Kim, J. Lee, and K.-C. Kwak, "Ensemble three-stream RGB-S deep neural network for human behavior recognition under intelligent home service robot environments," *IEEE Access*, vol. 9, pp. 73240–73250, 2021.

[4] I. Laptev, "On space-time interest points," *International Journal of Computer Vision*, vol. 64, no. 2, pp. 107–123, 2005.

[5] I. Kviatkovsky, A. Adam, and E. Rivlin, "Color invariants for person reidentification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1622–1634, 2012.

[6] R. V. H. M. Colque and C. Caetano, "Histograms of optical flow orientation and magnitude and entropy to detect anomalous events in videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 3, pp. 673–682, 2016.

[7] A. Wood, R. Mack, and M. Turner, "Developing self-determined motivation and performance with an elite athlete: integrating motivational interviewing with rational emotive behavior therapy," *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, vol. 38, no. 4, pp. 540–567, 2020.

 [8] T.-B. Bo, X.-Y. Zhang, K. D. Kohl, J. Wen, S.-J. Tian, and D.-H. Wang, "Coprophagy prevention alters microbiome, metabolism, neurochemistry, and cognitive behavior in a small mammal," *The ISME Journal*, vol. 14, no. 10, pp. 2625–2645, 2020.

 [9] J. W. Davis and S. R. Taylor, "Analysis and recognition of walking movements, object recognition supported by user interaction for service robots," *IEEE*, vol. 1, pp. 315–318, 2002.

[10] S. Aslani and H. Mahdavi-Nasab, "Optical flow based moving object detection and tracking for traffic surveillance," *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, vol. 7, no. 9, pp. 1252–1256, 2013.

[11] M. Blank, L. Gorelick, and E. Shechtman, "Actions as space-time shapes,"vol. 1, pp. 1395–1402, in *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)*, vol. 1, pp. 1395–1402, IEEE, Beijing, China, 2005.

[12] R. Ashraf, S. Afzal, A. U. Rehman et al., "Region-of-Interest based transfer learning assisted framework for skin cancer detection," *IEEE Access*, vol. 8, pp. 147858–147871, 2020.

[13] T. Donker, I. Cornelisz, C. Van Klaveren et al., "Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial," *JAMA Psychiatry*, vol. 76, no. 7, pp. 682–690, 2019.

[14] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial," *JMIR mental health*, vol. 4, no. 2, p. e19, 2017.

[15] A. Käll, S. Jägholm, H. Hesser et al., "Internet-based cognitive behavior therapy for loneliness: a pilot randomized controlled trial," *Behavior Therapy*, vol. 51, no. 1, pp. 54–68, 2020.

[16] J. H. Wright, M. Mishkind, T. D. Eells, and S. R. Chan, "Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety," *Current Psychiatry Reports*, vol. 21, no. 7, pp. 62–69, 2019.

[17] P. Lindner, "Better, virtually: the past, present, and future of virtual reality cognitive behavior therapy," *International Journal of Cognitive Therapy*, vol. 14, no. 1, pp. 23–46, 2021.

[18] W. S. Agras, "Cognitive behavior therapy for the eating disorders," *Psychiatric Clinics of North America*, vol. 42, no. 2, pp. 169–179, 2019.

[19] L. J. Kreuze, G. H. M. Pijnenborg, Y. B. de Jonge, and M. H. Nauta, "Cognitive-behavior therapy for children and adolescents with anxiety disorders: a meta-analysis of secondary outcomes," *Journal of Anxiety Disorders*, vol. 60, pp. 43–57, 2018.

[20] E. Forsell, M. Bendix, F. Holländare et al., "Internet delivered cognitive behavior therapy for antenatal depression: a randomised controlled trial," *Journal of Affective Disorders*, vol. 221, pp. 56–64, 2017.

[21] L. A. Brotto, S. Bergeron, B. Zdaniuk, and R Basson, "Mindfulness and cognitive behavior therapy for provoked vestibulodynia: mediators of treatment outcome and long-term effects," *Journal of Consulting and Clinical Psychology*, vol. 88, no. 1, pp. 48–64, 2020.

Retraction

Retracted: Internet False News Information Feature Extraction and Screening Based on 5G Internet of Things Combined with Passive RFID

Computational Intelligence and Neuroscience

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] D. Hu and Q. Xia, "Internet False News Information Feature Extraction and Screening Based on 5G Internet of Things Combined with Passive RFID," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9696472, 11 pages, 2021.

Research Article

Internet False News Information Feature Extraction and Screening Based on 5G Internet of Things Combined with Passive RFID

Dahai Hu and Qiong Xia [ORCID]

School of Journalism and Communication, Wuhan University, Wuhan 430072, China

Correspondence should be addressed to Qiong Xia; 00002737@whu.edu.cn

In this paper, the authenticity of news information on the 5G Internet of Things (IoT) is studied, and a network false news information screening platform is designed and optimized by IoT combined with passive RFID. The electronic license chain based on data sovereignty is established, in which, combined with the identity identification and strong correlation ability based on the electronic license chain, a cross-industry, cross-business, and cross-field behavior record base database is formed; then, a digital library is constructed based on this base library; finally, through data sharing and management, a false news information feature extraction and screening platform is formed for the orderly management and reasonable dispatch of government resources and reducing various risks. The main functional modules implemented by the platform are the acquisition of news data and comment data, the retrieval and analysis of news data, the false detection of online news, and the visualization of false news data. However, there is still much public who are not aware or do not understand that news truth is this dynamic form. Therefore, this paper aims to inform the public that news truth in news context is a dynamic process by 5G Internet of Things combined with passive RFID. The public understands the circumstances where news truth may be dynamic truth to avoid being misled by false news.

1. Introduction

News truth is the lifeline of news; the news, true or not, will affect the public's trust in news; therefore, the press always takes truth as the life of news and insists on news truth as one of the most basic business norms and the most important working principles [1]. Scholars in different periods will continue to change the research and results of news truth with the requirements of technology and society, constantly exploring and deepening that the connotation of news truth is the source of making news alive and everlasting. In recent years, news inaccuracies have frequently appeared, mostly in new media platforms, and the status and authority of news media have been greatly impacted, and the truthfulness of news has been questioned, so in this context, it is of practical significance to rethink the truthfulness of news reports in the context of news [2]. By reexamining the dynamic process of news authenticity, journalists, especially the audience, can understand that news authenticity is more prominent in the

new media context as a dynamic expression and realize that "news authenticity is a dynamic process" to reach a consensus and reshape the authority of news media, which is important to reduce false news and maintain social stability. And it is of great significance to reduce false news and maintain social stability [3].

In the network news release link, the media plays an important role, and creating a healthy network environment requires the media to supervise each checkpoint. Traditional media is strengthening integration and development, continuing to play the role of authoritative media in the Internet environment, using the Internet to present news content in traditional media to users in more diverse forms, but for now, traditional media has formed in the Internet of Things [4]. The influence of Internet media is growing. Therefore, the governance of Internet false news incidents still needs to rely on Internet media, by improving the media literacy of online media and exploring the characteristics of its gate-keeping mechanism, effectively preventing the spread of

online false news incidents [5]. News production in news contexts affects the production logic of news truth; news verification before reporting is converted to verification after reporting, and news truth is transformed from conclusion truth to process truth, all of which are closely linked to the characteristics of news contexts. The algorithmic mechanism is a new thing in the Internet era, and the research in this field in the academic industry is not yet mature. As the core mechanism of Internet information flow, machine algorithms have a certain impact on citizens and society through the collection and processing of information distribution [6]. The balance between technical rationality and value rationality can make technology better serve human beings, and the opposite is easy, to put the cart before the horse. How to balance the two requires contemporary researchers to propose more advanced and contemporary theories as to the basis. Most scholars have selected a definite ethical issue to study the phenomenon of this new phenomenon, but few scholars have looked at the essence of the phenomenon to analyze its intrinsic problems.

Network fake news is defined as false information posted by users on social media, that is, news information or rumors that are not consistent with the facts and cause a certain amount of public opinion released by individuals without official confirmation. Chapter 1: Introduction: this chapter first summarizes the research background and significance, research content, development trend, and application fields of Internet false news information feature extraction and screening based on 5G Internet of Things combined with passive RFID and finally gives a detailed description of the main research content of this article and the organization structure of the article. Chapter 2: Research on the Internet of Things News Screening Platform: the current research status is introduced. At the same time, an algorithm for feature extraction and screening of Internet false news information based on 5G Internet of Things combined with passive RFID is proposed, and an optimization platform for Internet false news information is established. Chapter 3: Result Analysis: with the optimization algorithm proposed in this paper as the core, the experimental process and experimental results are described and analyzed, and the effectiveness of the method proposed in this paper is verified through comparative experiments. Chapter 4: Conclusion: this chapter summarizes the full text, summarizes the main points of the full text, and looks forward to future work.

2. Research on 5G IoT News Feature Extraction Platform

2.1. Related Work. A study by Miloslavskaya pointed out that secure, well-attributed data sources are critical to data accountability, forensics, and privacy [7]. A decentralized trusted cloud data source architecture using IoT is proposed. IoT-based data sources provide tamper-proof records, increase the transparency of data accountability in the cloud, and contribute to the privacy and availability of source data [8]. In the paper, ProvChain is designed and implemented to collect and verify the architecture of cloud data sources by embedding source data into IoT transactions utilizing cloud

storage scenarios and selecting cloud files as data units to detect user actions to collect source data [9]. The results of the performance evaluation of this approach show that the source user privacy and reliability of cloud storage applications are minimal, so the source data collection, source data storage, and source data verification are needed [10]. The study by Hoang points out that data-based encryption, especially attribute-based encryption based on ciphertext policy, plays an important role in data sharing [11]. The study of an efficient and privacy-preserving traceable attribute-based encryption scheme for IoT is proposed. The scheme uses a preencryption technique to quickly generate ciphertexts and hides the attributes of anonymous access control structures by using attribute bloom filters. The source of misused keys can be audited in case of misuse.

Uslu et al. collected official rumor information on news platforms as an experimental data set. Through research, they found that the geographic location and terminal equipment used by users when posting Weibo have an important impact on rumor recognition. The feature is added to the feature set of the rumor, and the rumor detection based on the Weibo platform is realized by constructing an SVM classifier [12]. Kumar combined 5G IoT electronic data to become common evidence, from the perspective of judges reviewing the authenticity of electronic data, explored the possibility of electronic data being tampered within various aspects of collection, extraction, preservation, and transmission due to its intangibility and vulnerability to destruction, as well as the difficulty of discerning the authenticity and integrity of electronic data, which affects the admissibility and probative value of electronic data, and concluded IoT-based technology for multiparty deposition, to ensure the authenticity, security, and validity of electronic data conclusion [13]. Research on the compatibility of UHF RFID and cellular networks is based on reflection modulation and proposes a design scheme for passive cellular IoT systems, including cellular IoT base stations, passive terminal software and hardware architectures, and mainstream benefits compatibility. Although it is still a theoretical study, it can be seen that the research of cellular networks supporting passive IoT has also been put on the agenda, and the research direction of 5G supporting passive IoT access also has a certain foundation. In the past development process of the Internet of Things, the passive Internet of Things seems to have no intersection with the cellular network. Today, the vision of allowing 5G to carry hundreds of billions of Internet of Things access is indispensable for passive IoT support. Driven by this goal, cellular mobile communication networks are closely integrated with passive IoT for the first time and become the future largest group in 5G connections.

Fake news refers to news reports that do not reflect the objective truth and contain false elements and can be divided into two categories: one is news reports in which the communicator intentionally fabricates facts out of thin air for some purpose and disseminates them; the other is news reports that are generally true, but some details of the description appear to be false or intentionally or unintentionally misinterpreted or falsified. Reverse news is "news

that is initially reported in the opposite direction, and as the news report goes deeper and deeper, the truth is presented to the readers more objectively and comprehensively, and the readers' position is sharply reversed and they show the opposite attitude from before." Reversal news is distinct from fake news. This article combines news text similarity features, news and comment data emotional tendency features with some statistical-based user features to construct feature vectors, and uses IoT classifiers to train the data to obtain the final fake news detection model through the model "detect news data and display the classification results in the system."

2.2. Information Feature Extraction Algorithm. In intelligent algorithmic backend applications, the algorithm decides what content can enter the distribution channels, what information each user likes to read, and what news reports can occupy the main part of the distribution content and seem to evolve into the decision of the audience, which will lead to algorithmic news catering to the entertainment needs of the audience, ignoring the impact of the news itself on the public, industry, and society. Usually, post-truth news is the result of a significantly different public opinion sentiment that accompanies modern public opinion and is an important sign of the reversal of public opinion [14]. Fake news will not necessarily be clarified. The "reversal" of reversal news is the process by which the truth is continuously discovered. Reversal news will generally cause strong repercussions and create a strong public opinion situation.

Robots for data collection have become an important source of information for new media; a few days ago, the algorithm recommended news mainly relying on three production push mechanisms: one is the user portrait algorithm, in full accordance with the user hobby mapping recommended information, using the similarity of two or more user attributes to find the same cluster of users. These three content distribution models have in common the pushing model of giving different weights to algorithms to seek attention to the news, which makes it difficult for the public to scientifically select the social and cultural values embedded in the news, which requires gatekeepers to adhere to the confirmation and review of information, to guide the public to rational behavior, and to ensure that objective and accurate news will not be lost in the ocean of information. Therefore, in the stage of intelligent media, the existence of gatekeepers is reasonable, and the breakthrough can be made by personalized news, with algorithms and professional editors acting as gatekeepers, avoiding algorithms occupying the main position of gatekeepers, ensuring that news content meets the ethical and moral requirements of journalism, and reshaping the authority of gatekeepers, as shown in Figure 1.

The news media has the function of guiding social opinion and simultaneously setting important events in society as fixed issues. Although robots have changed the news production and content distribution mode, robots cannot replace the monitoring function [15]. The mass media in the traditional sense represents the public opinion

and is a collection of public opinions and issues. Public opinion supervision is completely different from rigid forms of supervision such as administrative supervision; it is a special form of soft supervision without strict regulations and requirements, and thus not compulsory, but it can play a role in educating and spreading mainstream values to the public. The media assumes the role of a public issue person in public opinion supervision, and as robotic writing becomes an important development trend in the media industry, it does not completely change the status of machine algorithms replacing human beings as the main body of supervision, which seems to be a cause for celebration. By continuously adding or deleting features, the accuracy of the model is tested on the verification set to find the optimal feature subset. The wrapped method usually has higher accuracy because of the direct participation of the model, but because each feature changes, the model must be retrained, so the calculation cost is high. Another disadvantage is that it is easy to overfat.

$$F(D_i) \leq H_i - \|T_i\| \|\nabla T X\|_i \|d_i N_i^2\|. \tag{1}$$

In the digital world, everyone has a physical presence, and all communication and collaboration will be based on the digital world as the norm and core, while the customary physical world becomes an aid and supplement to the digital world. This process is not only the realization of technology. The IoT security gateway reports news information recognizes news information in the unified management platform, and displays news distribution and true and false status in real time. Combined with the internal network isolation and baseline detection of the Internet of Things gateway, it blocks terminal intrusion detection and propagation.

$$X = \lim_{N \longrightarrow \infty} \sum_{i=1}^{N} \beta_i^n (d_i - 1). \tag{2}$$

We can achieve objective, fair, and scientific supervision and full supervision of credit evaluation through the comprehensive application of IoT and other technologies and, at the same time, strengthen the supervision of credit evaluation agencies to ensure that their evaluation is fair and untamable [16]. One is to use data sovereignty IoT technology to guarantee the operation right and basic bookkeeping right of IoT, and the core institution of data sovereignty to carry out the support and application of operation, to effectively guarantee the safety of people and national information property; the second is to use IoT technology to supervise the operation process of power operation, livelihood application, public service, and other affairs in the whole process.

To observe *X*, the signal *X* is projected using a matrix of *M* to obtain the observation vector *Y*.

$$Y = A_i^{CS} (\beta - 1) \beta_i^n. \tag{3}$$

IoT technologies are devoted to developing key technologies at the protocol and infrastructure layers that can take credit scoring-related data collaboration to a higher level. Meanwhile, a high-performance and secure credit data exchange network
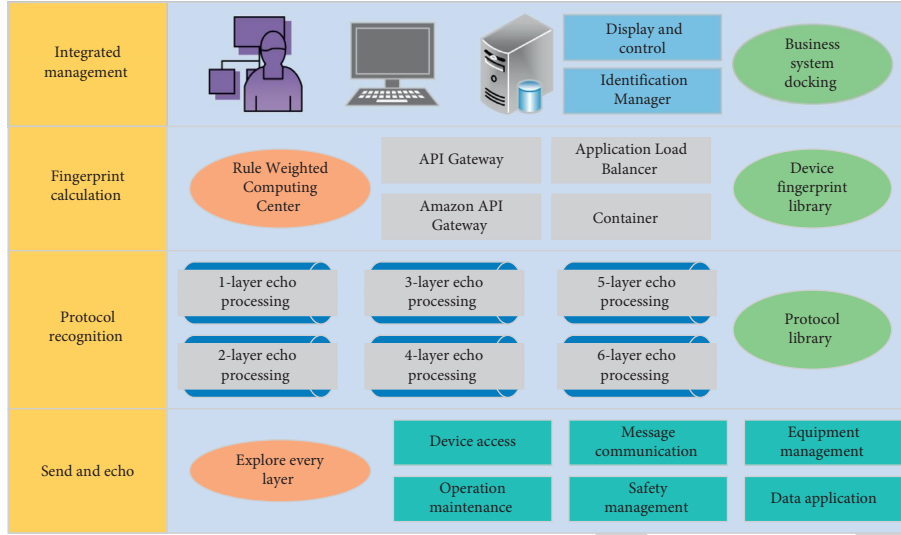
Figure 1: IoT information screening model architecture.

based on IoT technology can improve data quality, reduce data collection costs, expand data coverage, and ultimately create better credit scoring models through artificial intelligence used to update and calculate the network parameters that affect the model training and model output to approximate or reach the optimal value, thereby minimizing (or maximizing) the loss function. Therefore, IoT technology can not only play the advantages of high security and benefit-sharing, but also achieve the performance requirements needed for industry scenarios. Formula (4) is a performance statistical function, where β represents the statistical type, and n represents the statistical period.

$$\beta = \arg\max \left\| \beta_i^n \right\|^2. \tag{4}$$

Because laws and regulations can only be remedied after the fact through state intervention, while technical rules can be prevented beforehand through code. The function and role of law and code under the Internet society can be considered high and low. However, the translation of legal rules into technical rules is not an easy task. A legal code is a general rule that is essentially ambiguous and written in language; a technical code, in contrast to a legal code, can only be expressed through the code and necessarily relies on algorithmic forms and digital models. While code increasingly mimics or even replaces some of the traditional functions of law in the digital world, in the last few years, especially since the emergence of IoT technologies and smart contracts, the law has also come to display some of the characteristics of code. The feature extraction and screening of false news information are shown in formula (5), where U represents the feature set, t represents the collection time, s represents the screening method, and c represents the extraction method.

$$U_t = \frac{y_i^2(t,s)\left(1 + e^{x_m c_m \nabla A}\right)}{e^{x_m c_m \nabla A}}. \tag{5}$$

This information recommendation method helps users find the information they are interested in and read daily in a sea of information, reducing the time spent on information search, and people seem to have used to this way of receiving the news [17, 18]. The algorithm replaces the user's way to find information and invariably replaces the individual user's information selection, creating private customization of the information content platform, allowing irrelevant news to be lost in the sea of information, and making the user refuse to receive social issues from other public areas, which will generally reduce the public's social cognitive ability in the long run.

2.3. Fake News Optimization Combined with Passive RFID. It is reflected in the internal verification process of journalists, which is the most reflective link of the main position of journalists and belongs to the prereporting verification. For a long time, journalism at home and abroad has always held "news truth" as the lifeline of news and requires journalists in the process of news reporting must ensure that news events are reported in a completely true and objective way; therefore, journalists in the acquisition of newsworthy information about recent events must go through a verification [19]. In the process of verification, journalists must collect as many accounts as possible from different parties or eyewitnesses to restore and reproduce the actual event from multiple angles and in a comprehensive manner. Verification of news content is a necessary process to ensure that news reports are authentic and credible, and so is news production. Figure 2 represents the optimized architecture of the news information screening platform combined with passive RFID. The main purpose of the online fake news detection system is to detect online fake news. The system contains four main functional modules, namely, the acquisition module of Internet of Things network news data and comment data, the analysis and retrieval module of network news data, and the network false news detection module, and the data analysis result visualization module, among which the IoT false news detection module is the most important functional module in the system.
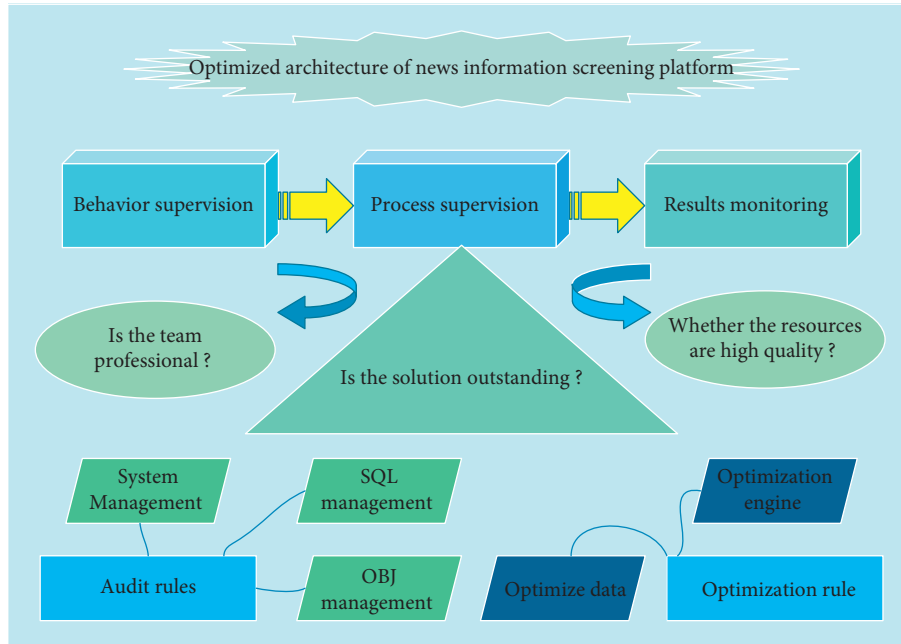
Figure 2: Optimized architecture of news information screening combined with passive RFID.

The process of authenticity identification service design is as follows: firstly, the client node in the IoT platform calculates the digital digest of the local file, calls the authenticity identification API, and initiates a request to the server node; then the server node authenticates the network certificate held by the node, and if the verification passes, the node is authenticated, and the file access authority is verified; if the verification fails, an error message is returned, and the call is terminated [20]. If the verification passes, and the nodes on the chain reach consensus, the relevant data in the process will be written to the block and packaged, and the corresponding message will be returned to the client node; otherwise, an error message will be returned, and the call will be terminated, thus ending the authenticity identification service.

News reporting is the reporting of newly occurring or newly discovered facts, and the object of its reporting is the matter that is available in the objective world. Epistemology believes that the objective material world is always in the process of movement and change; therefore, from the perspective of the object to explore the dynamic process of news reality, it is necessary to clarify a point of view; that is, the fact itself is in a state of motion of constant change and development, and then the news reality corresponding to the event will also necessarily be in a dynamic process [21]. Many news reports of the event with special characteristics and complexity, in those instant reports of the news process, due to the reporting of the event at the time of some matters or details have not been fully revealed; instant reporting can only present the facts that have been revealed partly, and then with the development of the event for follow-up reports, in such a continuous and dynamic reporting, the overall reality of the news event also shows a dynamic presentation. In such continuous and dynamic reporting, the overall truth of the news event is dynamically presented. Second, the fragmented news text in the news context can hardly present the complicated and objective world completely. Therefore, the news truth from the object perspective is also in a dynamic development process with the development of news events.

Traditional mainstream media are limited by technology to a longer time limit for news reports and limited channel resources, and news reports are more inclined to one-time final reports, with many news being reported only after the development of the matter is over, and after in-depth digging and verification by reporters. Therefore, the content of these reports is completed, and the news reality shown to the audience is also more complete. Only a small number of news events with a relatively large impact, such as sudden disaster events, are released for the first time to inform the public, and then follow-up reports of the events are released continuously. In contrast, news reporting in the news communication context has shown recent changes and forms due to the real-time nature of news releases and the brevity of news content, as well as the competitive pressure of information release channels. To grab headlines, professional news media often release concise news stories to let audiences know that something has happened before it is finished and then follow up on the news to grab headlines. The complex communication ecology of the news communication context drives news reports to keep pace with the development of news events and makes the dynamic characteristics of news truth more prominent. In the news communication context, news truth also presents a gradual and dynamic process because of the development process of news events.

IoT technology has the characteristic of "data sharing," which refers to the sharing of data storage and exchange level

[22–25]. The IoT network makes multiple copies of the data on the chain and stores them on multiple storage nodes at the same time, so from the data point of view, the data is shared to multiple nodes and can be downloaded from any node. However, from the perspective of data content, the platform can encrypt the data content through encryption technology. Even if we get the encrypted data, we cannot access the actual information contained in the data without the decryption key. Therefore, the platform identifies the legitimate users based on the account mechanism, assigns the users the contents they have a right to view and the decryption key through the permission mechanism, and ensures data security by encrypting the data on the chain through the encryption technology. The platform supports the information security mechanism based on permission and encryption functions. In simple terms, the integrity file platform needs to be completed with a system that supports identity-based permission control and password mechanism and only responds to user requests through the IoT gateway, specifically including filtering out requests from illegal users and cross-authority requests from legitimate users.

All data compiled in this study included all online news headlines from seven major websites with names in the title and time from 2010 to 2019. A total of 8825000 news headlines were collected, and the results are shown in Table 1.

The network false news detection module detects false news based on the constructed test model. In the system design of this article, the test model generated in advance based on the training data is stored in a txt file, and the news collected in this system can be detected and identified. When you click on fake news analysis, you can use a visual mode to show the trend changes in the number of comments related to the news and the analysis results of the text and comments.

3. Analysis of Results

3.1. News Information Feature Extraction Results. Social platforms such as microblogs, which are based on Internet technology, have the functions of both information dissemination and social interaction, allowing users to express their opinions or respond to others' opinions through comments, retweets, and private messages. The dissemination of information on the microblogging platform is not restricted by time and space, and users can spontaneously and actively participate in discussions, actively verify the information, and provide other information after news reports are released, which invariably constitutes a unique "self-clarification" mechanism in the online environment, as shown in Figure 3. Among them, the information of the day accounted for the highest proportion, reaching 32.5%, and the proportion of information over two weeks was the lowest, accounting for only 11.2%.

Among them, the information of the day accounted for the highest proportion, reaching 32.5%, and the proportion of information over two weeks was the lowest, accounting for only 11.2%. Those that are longer than two weeks are the event itself that is still in the development process, so it takes

TABLE 1: Results of the collected news volume.

Year	Value
2010	779854
2011	805974
2012	1685748
2013	1694715
2014	1005895
2015	1234874
2016	1654287
2017	1029574
2018	900125
2019	958426
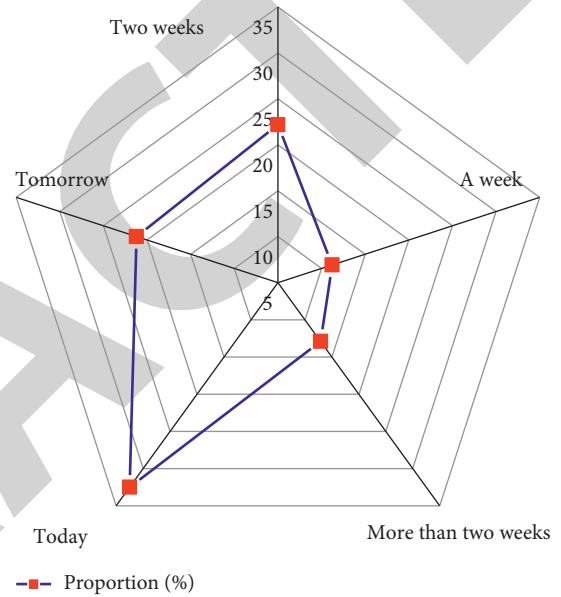
FIGURE 3: Length of news true clarification.

a relatively long time. Overall, the real clarification time of news in the context of Weibo is relatively short, and clarification can be obtained relatively quickly.

Filter out the useless comment data, and extract the features of Weibo content and comment-related content, based on the original feature, the combination of the original feature and the newly extracted text feature, and the original feature and the newly extracted text feature. The combination of text and comment features was experimented separately. Through the data in the test, the reliability and performance stability of the IoT platform are reported. However, in the process of false analysis of news and sentiment analysis, due to the processing of more data, there will occasionally be longer response times. The detection efficiency of the IoT platform still needs to be improved.

Through the analysis of 60 cases and overall, it is found that 73.33% of the news reports present dynamic reality because of the lack of verification or being reported and reprinted without verification, 11.67% of the news reports are fictitious news to win the audience's attention, and other factors include poor professionalism of journalists, unfinished events, and lack of authoritative information, as shown in Figure 4.

Figure 4 reveals that the main reason for the prominence of the dynamic process of news in the current microblogging context lies in the lack of prerelease verification of news reports. The cause for the widespread of many fake news stories is that professional news media simply copy and paste them without verifying the news content, all at the cost of losing the credibility and authority of the news. The phenomenon of backstage news foregrounding has emerged along with the weakening of the verification process in the professional news media. Goffman introduced theatrical elements and staging techniques into interpersonal interactions and put forward "drama theory," which compares interpersonal interactions to personal performances and proposes the concepts of front stage and backstage. The front stage performance is more formal, while the backstage performance is relatively relaxed and unruly. Some studies have introduced Goffman's drama theory into the news reporting industry and explored the relationship between news media and audiences, arguing that news also has a front and backstage, in which the production work before the release of news reports is the backstage, TV, newspapers, and official microblogs belong to the front stage, and each news report is treated as a front stage performance, and the news reports that audiences are exposed to are front stage acts, and the previous news front and backstage boundaries and behavioral segmentation. The previous news front and backstage boundaries and behaviors are relatively clear. However, in the microblogging context, the front and backstage boundaries of news reports are blurred, because the verification process of news media is weakened, which reduces the authenticity of news reports, and many news are reported after netizens raise questions, or some knowledgeable people point out that the news reports are not in line with the actual situation, which leads to the verification behavior of the backstage being moved to the front stage and the front stage turning into the backstage. This is one of the reasons why news reality in the microblogging context presents a dynamic process.

Figure 5 shows the changing pattern of network size at each dissemination stage, and the numbers of nodes and edges of the network are the indicators of network size. Among the four stages of online opinion dissemination, the numbers of nodes and edges are the largest in the outbreak period with 3864 and 1610, respectively, indicating that the network scale is the largest at this time, and most microblog users are involved in opinion dissemination. In the latent stage, the network size is the smallest, and the numbers of nodes and edges are small, but the difference is not very large, while, in the growth and decay stages, the number of nodes of the network is greater than 1000, and the number of edges is about 330, indicating that although many users participate in the discussion in these two stages, only a small number of users have retweeting relationships with each other.

From the trend of the maximum node degree of network nodes in Figure 5, it can also be seen that although the maximum node degree of network nodes differs significantly from the network nodes and edges in terms of number, the trend of change does approximate the same. The maximum

node degree of network nodes also reaches its maximum value during the outbreak of public opinion. News from different information sources is brought together, and news truth becomes more complex and difficult to discern. Often, the public must increase their capacity to select and screen information and consciously develop their thinking about news counterfeiting. Do not jump to conclusions blindly, but keep a skeptical attitude. If necessary, you should combine life experience and common sense to make judgments, just like in daily life, and not believe in things that fall from the sky; similarly, in the microblogging platform, you should not believe in information that is inconsistent in wording and things that do not make sense and also be easily convinced and spread it, but you should reserve your opinion about some uncertain information. It is better to wait patiently for the professional media to provide authoritative information, to minimize the negative impact brought by the process of real news dynamics.

3.2. Platform Optimization Results. As shown in Figure 6, the latent period of public opinion has both large network density and average network degree, but the average shortest distance length of the network in this period is particularly large, which makes the connection between users in the latent period very close but hinders the dissemination of public opinion information more difficult due to the longer dissemination path. While the values of network density in both the outbreak and decay periods of public opinion are relatively small, and information transmission is not very easy, the average shortest distance length of the network is much smaller than the rest of H periods, and the average degree of the network is also relatively large, so that, in the outbreak period, public opinion information is transmitted quickly among users, and communication among users is more frequent; on the contrary, in the decay period, the average shortest distance of the network is very large, and the average degree of the network is relatively small, so that, in the decay period, public opinion information is spread slowly, and users are less connected to each other and interact less.

Figure 7 shows the statistics of the number of news retweets. These users have a common characteristic; that is, they have a certain status of speech in the society, and the blog posts and opinions they publish will be responded to by a group of a not small number of people and exert a large influence, which is what we call opinion leaders. These users have a huge circle of fans and a large influence, so they do not have to deliberately publish numerous blog posts or follow others on purpose. When public opinion occurs, the information released by each person only represents the opinion, and the quality of information from certified users is relatively much higher, which also plays a role in the supervision of online information. These users will post or forward blog posts that can represent their emotional tendencies, thus influencing the remarks and emotions of other information recipients and playing a role in the spread of public opinion.

Figure 8 indicates the statistical information of the system node data. The system performance and system

A: Unverified   D: Missing information
B: Fictitious content   E: The event is not over
C: Lack of expertise
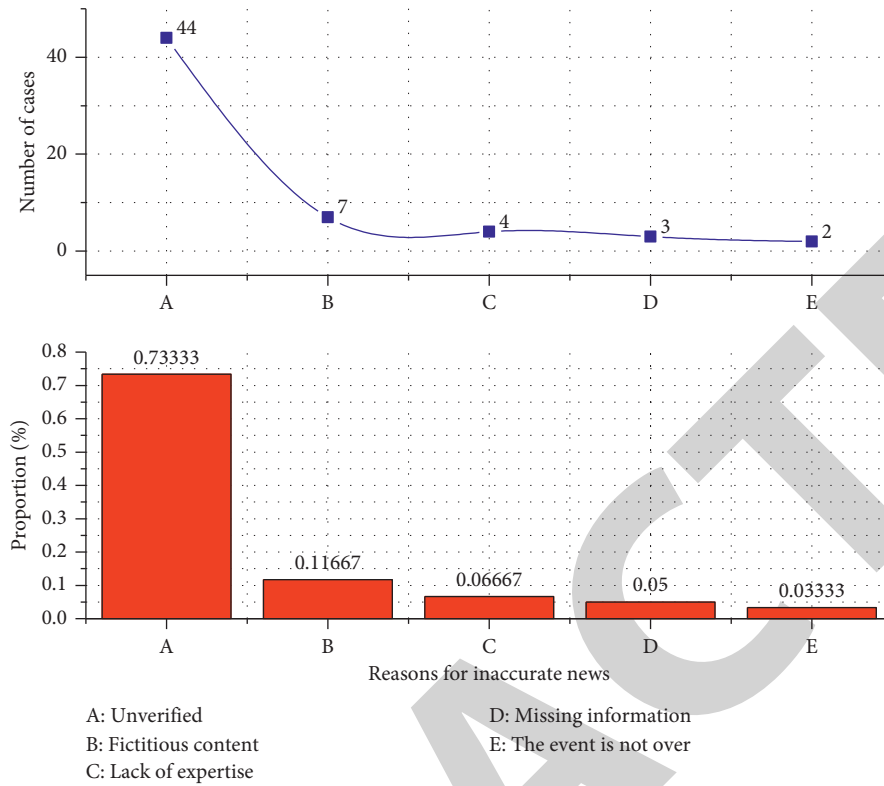
Figure 4: Reasons for news misrepresentation.
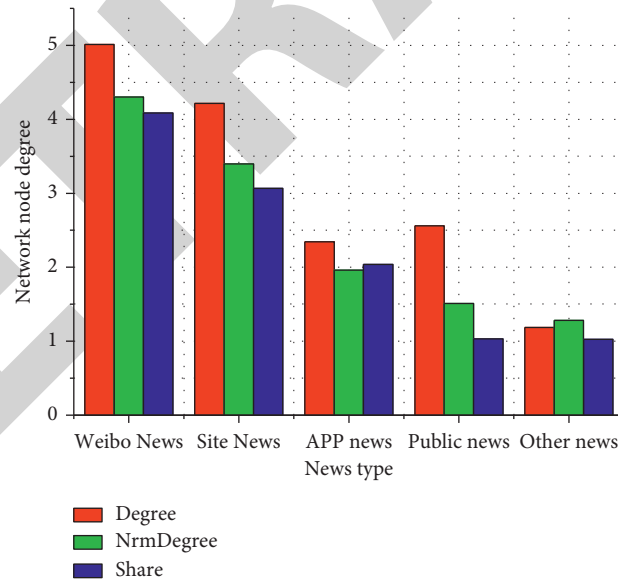
Figure 5: Network node degree during the decay.

operation data of all aspects of the constructed integrity file platform meet the technical requirements, realize the electronic license chain based on data sovereignty IoT technology, and can realize the efficient application of electronic license in anticounterfeiting, impersonation verification, privacy protection, safe sharing, visualization, and authority. This set of integrity file systems built based on integrity data sources with communal sharing in

the whole field has laid a solid foundation for the subsequent in-depth research of data sovereignty IoT technology.

In the process of modern news dissemination, the report form of any piece of news has to go through many processing processes. In these processing processes, news facts are recognized, formed into images, latent processing, written expressions, value judgments, and
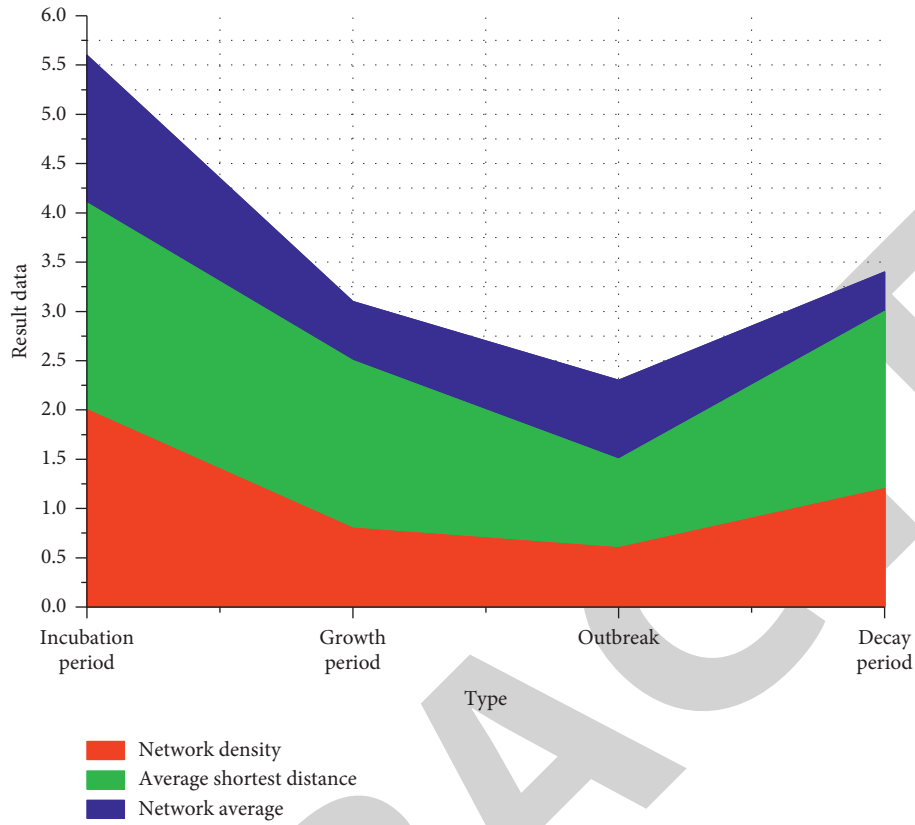
Figure 6: Trends in network characteristics at each dissemination stage.
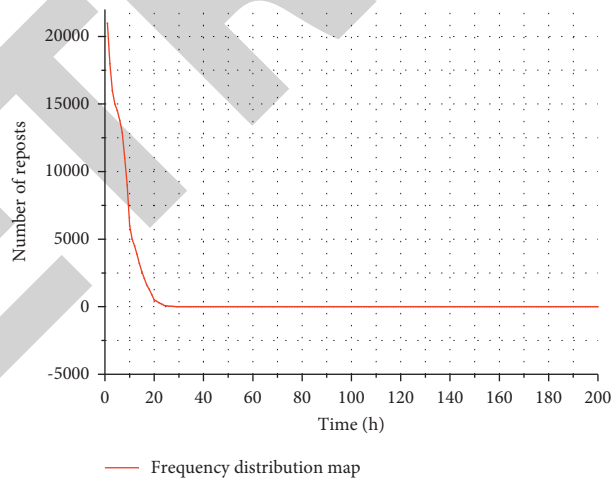
Figure 7: Distribution of the number of forwarding.

disclosures. There are many links such as publication, and each link is operated by humans. This not only helps ensure the authenticity of the news, but also greatly helps the news writing process and enrichment of the content. It is also the basis for writing good articles. As journalists in an era of unprecedented fierce competition, we should remember that any report we make must be personally and meticulously investigated and researched before it can be published responsibly to a broad audience. As long as this is achieved, the authenticity of the news will be fundamentally guaranteed, and the lifeline of the news media will be fundamentally ensured.
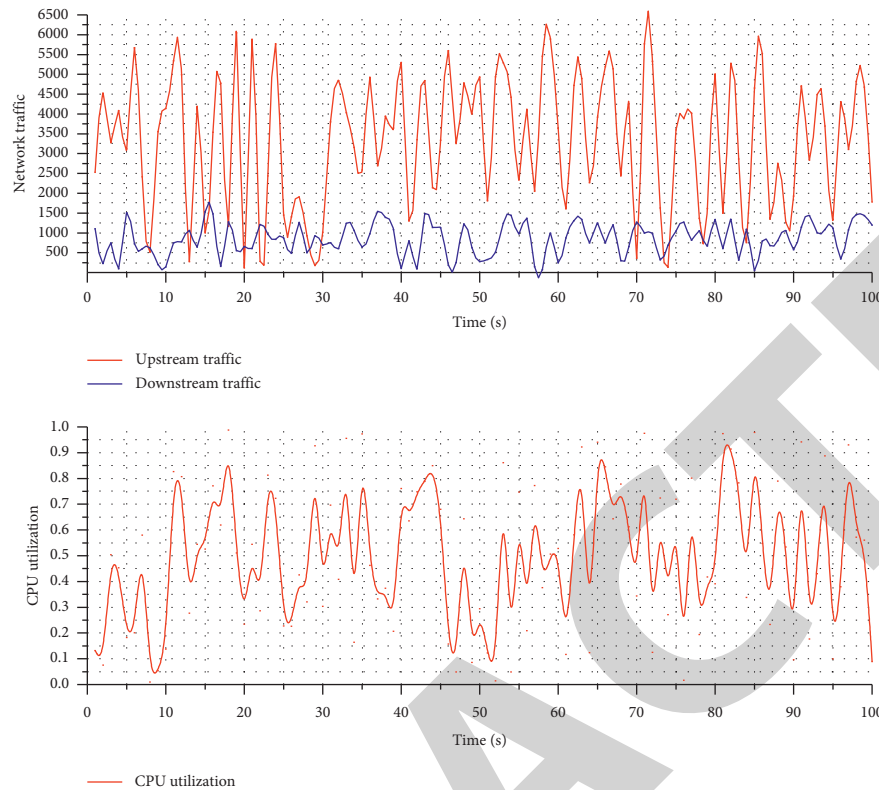
FIGURE 8: Node data.

4. Conclusion

The paper elaborates the theoretical system, principle, and development process of the key technology of the news information feature extraction platform based on data sovereign 5G Internet of Things and demonstrates and tests the prototype system of the news information screening platform combined with passive RFID. The truth of news reports is as a small part reflecting the overall social reality, and the facts of news reports from different angles and positions that finally present the real picture of the whole society through the stage and dynamic reports. Also, because the production process was more rigorous and thorough during the traditional media era, the truthfulness of the news was better assured. The limited human knowledge of technology and self-knowledge, coupled with the public's blind pursuit of the excitement and novelty of new technologies and the lack of corresponding rational thinking, has exacerbated the degree of ethical misconduct in journalism on some levels. Data sources can be expanded in future research. How to obtain these data from the media and how to analyze fake news purely from a textual perspective are a long-term and meaningful thing, so the screening of fake news still needs to be further investigated. The experiment and analysis of data mining technology are used to identify fake news and thus develop a more accurate fake news screening system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] E. Keane, K. Zvarikova, and Z. Rowland, "Cognitive automation, big data-driven manufacturing, and sustainable industrial value creation in internet of things-based real-time production logistics," *Economics, Management, and Financial Markets*, vol. 15, no. 4, pp. 39–48, 2020.

[2] S. AlZu'bi, B. Hawashin, and M. Mujahed, "An efficient employment of internet of multimedia things in smart and future agriculture," *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29581–29605, 2019.

[3] Y. R. Shrestha, S. M. Ben-Menahem, and G. Von Krogh, "Organizational decision-making structures in the age of artificial intelligence," *California Management Review*, vol. 61, no. 4, pp. 66–83, 2019.

[4] F. J. Zuiderveen Borgesius, "Strengthening legal protection against discrimination by algorithms and artificial intelligence," *International Journal of Human Rights*, vol. 24, no. 10, pp. 1572–1593, 2020.

[5] K. E. Jeon, J. She, P. Soonsawad, and P. C. Ng, "BLE beacons for internet of things applications: survey, challenges, and opportunities," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 811–828, 2018.

[6] S. N. Matheu, J. L. Hernández-Ramos, A. F. Skarmeta, and B. Gianmarco, "A survey of cybersecurity certification for the internet of things," *ACM Computing Surveys*, vol. 53, no. 6, pp. 1–36, 2020.

Research Article

Optimization of STP Innovation Management Mechanisms Driven by Advanced Evolutionary IoT Arithmetic

Tianxiang Wang,[1] Qingqing Ma [iD],[2] and Jinxi Li[3]

[1]*Faculty of Management and Economics, Kunming University of Science and Technology, Kunming, Yunnan 650500, China*
[2]*Anqing Normal University, Anqing, Anhui 246133, China*
[3]*Party School of the CPC Xinyang Municipal Committee, Xinyang, Henan 464000, China*

Correspondence should be addressed to Qingqing Ma; qingqing237877@aqnu.edu.cn

Since industrialization, manufacturing has been an important pillar of a country's economic development. Under the dual pressure of the new trend of global manufacturing development and the loss of competitive advantage of manufacturing industry, it is especially important to accelerate the enhancement of national high technology innovation capacity and the optimization of high technology policy innovation management mechanism driven by advanced evolutionary Internet of Things (IoT) arithmetic. The main of this paper thus introduces the effective method of optimization of high technology policy innovation management mechanism driven by advanced evolutionary IoT arithmetic. To study the optimization of high technology policy innovation management mechanism, a conceptual analysis of currently popular information technologies, such as big data technologies, artificial intelligence technologies, and Internet of Things technologies, and an overview of the application of these technologies in microgrids are given. In the paper, all factors are studied using the STP innovation management mechanism-based approach, and finally, all factors are classified into two categories of cause and effect factors by this approach, and the importance of all factors is ranked. Secondly, a wind power prediction algorithm based on data mining technology and an improved algorithm and a PV power prediction algorithm based on a deep neural network were established with the technical support of high-tech information technology such as big data and artificial intelligence. Finally, the majorization of high technology policy innovation management mechanism driven by advanced evolutionary IoT arithmetic is proposed.

1. Introduction

High technology innovation is the fundamental driving force of a country's development, and the high technology policy and innovation management mechanism of every country in the world today cannot hold a laissez-faire attitude towards high technology innovation in the process of economic development, especially in developed countries paying more attention to the development of high technology innovation [1], and their high technology policy and innovation management mechanism have been formulating a large number of high technology innovation policies to guarantee and promote the high technology innovation activities. Especially in the twenty-first century, mankind has entered the era of economic globalization, informatization, and networking, and the traditional economic development model has undergone significant changes. From the international perspective [2], an innovative economy will gradually become the mainstream form of economic development, innovation has become the focus of the development strategies of developed countries and regions, and innovative countries have become an important symbol of scientific and technological power. To gain long-term development advantages in the competition of economic globalization, scientific and technological innovation is imminent [3].

High technology is the source of a country's sustainable development, and high technology innovation is gaining

ground in economic and social development, becoming a core element affecting national competitiveness. Improving the level of national high technology innovation has become the focus of attention of scholars and innovation management mechanisms of high technology policies in various countries [4]. For a long time, the leading countries in high technology have taken enterprises as the main body of high technology innovation, but the management mechanism of high technology policy innovation plays an important role in the process of promoting high technology progress. Because of the high risk and uncertainty of both market and technology in the process of development, and the monopoly of enterprises on scientific and technological achievements, all these situations will lead to market failure in the process of allocation of scientific and technological resources. Therefore, the role of the innovation management mechanism of high technology policy driven by advanced evolutionary IoT calculus is revealed. Innovation supports development and high technology lead the future [5].

Through the analysis of high technology policy innovation management mechanism high technology innovation policy orientation research, it can be seen that, in the rapid development of high technology today, to improve productivity and promote social progress at the same time, the need for continuous innovation, how to properly guide the innovation of high technology, and by whom to guide has become the current concern. The purpose of this paper is very clear for this topic, aiming to study the role played by high technology policy innovation management mechanism in the process of high technology innovation and how high technology policy innovation management mechanism promotes the formulation of high technology innovation policy. Optimize the index system of high technology innovation policy assessment. Clarifying the content of high technology innovation policy, exploring the main factors affecting the effect of high technology innovation policy, and constructing a reasonable and effective assessment index system are the prerequisites for correctly assessing the effect of high technology innovation policy. Based on the connotation and composition of high technology innovation policy, this paper selects indicators from three aspects, such as high technology innovation foundation and input, output and result, and regional economic benefit, and adopts factor analysis to screen the indicators, and finally determines the regional high technology innovation policy assessment indicator system, which is a useful supplement to the construction of high technology innovation policy effect assessment indicator system and provides a reference for the future high technology innovation policy effect assessment indicator system. It is a useful supplement to the construction of the STI policy effect assessment index system and provides a reference for the optimization of the future STI policy effect assessment index system.

2. Relevant Studies

After the introduction of "innovation" into economic theory for the first time in the literature, Western economists have explored the issue of innovation in greater depth, mainly

from two perspectives: one is to introduce technological innovation as an exogenous variable into the production function to explain the role of technological innovation on economic growth; the other is to endogenize technological innovation [6]. It is believed that sustained economic growth does not rely on exogenous push, but endogenous technological progress. After the 1980s, a considerable number of studies on high technology innovation have been conducted, and their research is mainly based on the deepening and expansion of the endogenous economic growth theory, which contains organizational innovation, institutional innovation, and management innovation in addition to technological innovation. The literature studies innovation based on the management perspective and argues that energy creation is not enough to rely only on technological improvements, but also on other potential behaviors to improve the wealth creation of resources [7], such as management innovation. The literature includes institutional innovation as an endogenous variable in the innovation process and the empirical analysis of maritime transport productivity over the period 1600–1850 concludes that institutional improvements are more likely to cause productivity growth than technological changes [8]. Thus the literature argues that institutional innovation determines technological innovation rather than technological innovation determining institutional innovation. The literature investigates the role of income disparity in influencing the evolutionary path of innovative products and the returns to innovation, pointing out that the smaller the income disparity, the more conducive it is to stimulate innovative behavior, thereby promoting economic growth; conversely, it discourages innovative behavior and inhibits economic growth. The literature proposes the "quality of innovation," which further elevates the previous criteria for judging innovation, such as novelty [9], unconventionality, and creativity, to the level of quality, i.e. standardization, systematization, and low variance. The literature argues that a firm's innovation capabilities include R&D, production, learning, marketing, resource development, organization, and strategic management. The literature explores the relationship between the price of innovative products and innovation behavior, stating that the price of innovative products is not an exogenous variable and that increasing the price of innovative products can increase the profitability of innovation when the income gap widens; on the other hand, a price increase can lead to a decrease in the rate of market demand, which in turn can reduce the return to innovation. The literature argues that innovation is a management process and that the scientific application of methods, rules, and institutions can be the efficient rate of completing innovative work [10].

The literature defines technological innovation based on the continuum of innovative behavior, design, manufacturing, and commercial activities involved in the process of producing new products, using new processes or new production equipment, and this activity process is in the order of process innovation, product innovation, and innovation diffusion [11]. The literature analyzes innovation from the perspective of the main position of enterprises and

combines technological innovation with enterprises. It believes that innovation is a series of activities in which enterprises reallocate and combine production factors and conditions to obtain potential profit opportunities, rapidly capture market information, and establish highly efficient, effective, and low-cost production systems. In addition, the literature points out that technological innovation in a broad sense refers to the entire process from the conception of a new product or process to its development [12], technology diffusion, and realization of its market value. The literature suggests that technological innovation can be broadly defined as all actions from the conception of a new idea to the production of a new product, including not only the outcome of the innovation itself but also the introduction, diffusion, and application of the technology. The literature constructs an index system for evaluating scientific and technological capability from three perspectives: scientific and technological research and development capability, scientific and technological achievement transfer capability, and scientific and technological support capability, and conducts a longitudinal and cross-sectional comparative analysis of the scientific and technological innovation capability of City A [13].

Based on the endogenous economic growth theory, the literature introduces the endogenous knowledge stock into economic growth and measures the output of high technology activities by using high technology funding input, knowledge stock input, capital input and labor input, and human capital input as input indicators and explores the relationship and development trend between high technology input and high technology output performance in China. Based on the Cobb-Douglas production letter model with GDP as the dependent variable and labor input, capital input, and rate of scientific and technological progress as independent variables, the literature uses an econometric model to study the impact of scientific and technological progress on economic growth in country A during 1980–2004 and applies Solow's surplus to calculate the degree of contribution of scientific and technological progress to study the development of scientific and technological progress [14]. The literature argues that high technology innovation and economic development are complementary to each other, one is that high technology innovation drives rapid economic development, and the other is that economic development, in turn, provides strong support for high technology innovation. The literature adopts T-shaped correlation analysis to empirically study the contribution of economic investment in high technology and investment in scientific and technological talents in economic growth in country A during 2000–2007, and the results show that R&D investment has a very significant role in economic growth, while the role of scientific and technological talents in economic growth is not obvious. Economic growth is more dependent on the introduction of new technologies. In exploring the interaction between regional marine high technology innovation and blue economic growth, the literature points out to regional marine [15].

There is a high correlation of the growth of the regional blue-blue economy of ocean high technology innovation,

and the relationship between them is not simply linear; it is necessary to establish a synergistic and efficient interactive operation mechanism to guarantee and support the positive interaction between the two and promote the coordinated development of regional marine high technology innovation and blue economy [16]. The literature discusses the mechanism of the role between high technology investment, high technology innovation, and regional economic development and proposes that there is a causal relationship between the three; i.e. high technology investment supports high technology innovation, and high technology innovation and regional economic development support each other, and high technology investment and regional economic growth are mutually reinforcing [17].

3. Optimization of STP Innovation Management Mechanisms Driven by Advanced Evolutionary IoT Arithmetic

3.1. Advanced Evolutionary IoT Arithmetic Example. The concept of the Internet of Things (IoT) was first proposed in 1991, and since then IoT has gradually entered people's vision. The Internet of Things uses Radio Frequency Identification (RFID), wireless communication, and other technologies to construct a physical Internet covering everything in the world based on the computer Internet. In the future, the continuous breakthrough of modern information technology will greatly promote the development of the Internet of Things, with the "Internet + Internet of Things," "artificial intelligence + Internet of Things," and other concepts of the perfect combination [18].

The Internet of Things (IoT) is a kind of convergence application and enhancement after the development of modern information technology to a certain stage, based on RFID identification technology, to realize people and things, things and things can "communicate" with each other, forming an intelligent network that can sense the real world. In simple terms, it is mainly composed of two parts: the radio frequency tag and the decoder. The tag is attached to the surface of the item to be identified, and after entering the magnetic field, it receives the signal from the decoder and sends out the product information stored in the chip (passive tag or passive tag) by relying on the energy obtained from the induction current, or the tag actively sends out a signal of a certain frequency (active tag or active tag), and the decoder reads the information and decodes it and sends it to the central information system and finally performs relevant data processing and analysis. According to the three main aspects of IoT data processing, its logical functions are recognized as three levels, i.e., perception layer, network layer, and application layer. The perception layer mainly collects, identifies, and intelligently controls physical information in the real world; the network layer serves as a link between the perception layer and the application layer to guarantee the transmission, routing, and control between IoT information data; the application layer is the topmost layer in the three-layer structure, and its core functions include two aspects, one is to complete into the processing

and management of various data, and secondly is the integration of these data with various industrial applications to achieve real-time control of the physical world. The general framework of the IoT is shown in Figure 1.

The current IoT technology mainly includes key technologies in the fields of perception and identification, intelligent control [19], embedded system, power supply, and energy storage, new material technology, wireless communication, chip design, and manufacturing, etc. The breadth of the involved fields and the complexity of the involved arts make it an extremely huge technical project. According to the presidential framework of IoT, the IoT technology system can be divided into three levels, such as perception layer technology, network layer technology, and application layer technology. Its formula principle lies in

$$\int \sqrt{d^2 + x^2} = \frac{d\sqrt{b^2 - 4ac}}{2}. \tag{1}$$

The first formula of the article summarizes the algorithm carried out in the article and summarizes the essence of the article. It is the basis of the algorithm mentioned in the article, and the following formulas are also mentioned. Among them, the sensing layer technology is the basis of the whole technical framework, the network layer technology is the technical guarantee of the whole information transmission, and the application layer technology is the terminal of the input and output control. The IoT technology system is shown in Figure 2.

The core technology of IoT is quite complex and fine, while the current overall R&D capability of China's IoT technology is weak, and when facing complex application requirements, it will be found that there is still a gap between the chip technology, hub components, and other high-precision technologies and the international leading level. Therefore, the development of IoT technology needs to be combined with China's specific national conditions and the current situation of industrial development to give priority to the development of IoT technology and provide a scientific basis for the country's macrodecision. At the same time, it is necessary to accelerate the development of the standardization of IoT technology [20] and steadily promote the promotion and application of IoT, to improve the overall level of IoT technology in China. With the rapid development of information technology, the scale of the IoT market will gradually expand, not only bringing huge economic benefits to society but also effectively safeguarding national security. IoT technology can be widely used in basic strategic areas such as security, electricity, food, oil, etc. It can also prevent the destruction of strategic resources such as strategic bases, public facilities, and forest land by hostile forces, terrorism, and natural disasters. The extensive use of IoT technology at the national strategic level provides a strong technical guarantee for securing homeland security. Its formula lies in

$$\int \sqrt{d^2 + x^2} = \frac{1}{2} \frac{n!}{r!(n-r)!} + \sqrt{b^2 - 4ac}. \tag{2}$$

The Internet of Things (IoT) applies a variety of modern information technologies to various industries, which itself is a complement and enhancement of the integration of two, which is also a component and application area of IoT technology. The concept of IoT is the integration of IT technology into automation control systems to form an intelligent network system that is efficient, safe, and energy-efficient. Using various advanced technologies such as perception, modern communication, artificial intelligence, and automation, the IoT reorganizes production factors and supply chains extensively and becomes a realistic carrier of informatization-driven industrialization. Therefore, the development of the Advanced Internet of Things Safety is a natural combination of industrialization and informatization, a powerful tool and substantial driving force to promote the integration of the two, and a realistic need to accelerate the integration of the two, whose realization model is shown in Figure 3.

3.2. Optimization of High Technology Policy Innovation Management Mechanism.

High technology policy is the policy adopted by the government to promote the development of high technology and to use high technology to achieve national goals. Its policy texts are generally expressed in the form of circulars, plans, opinions, regulations, and laws and regulations. The theoretical basis of high technology policy includes three aspects, such as market failure theory, system failure theory, and scientific social contract theory. The objective of high technology policy is to promote economic and social development, the development of political civilization and national defense as well as the development of high technology, and the improvement of the country's comprehensive national power and international competitiveness. The high technology policy also contains the high technology talent policy in terms of the development of high technology talent. For the meaning of high technology policy, most scholars believe that high technology talent policy is a series of strategies, laws, regulations, decrees, ordinances, measures, and methods for managing and regulating the behavior of high technology talents. The high technology management system is the overall institutional arrangement of a country's high technology management, including the setting and management of high technology-related institutions, the process of high technology policy formulation and implementation, and many other elements, which plays an irreplaceable role in releasing national innovation vitality, improving the efficiency of the national research system, and maintaining the long-term stability of innovation.

There are several types of national S&T management systems: centralized, dualistic, and decentralized. Centralized systems usually involve the establishment of national ministries of high technology management at the national level to unify the country's high technology activities, and such unification usually facilitates the full protection and nurturing of research institutions and the conduct of long-term, high-risk research. However, scientific and technological development can easily run counter to actual needs or be distant from the mission of other government agencies, affecting the utilization of scientific and technological
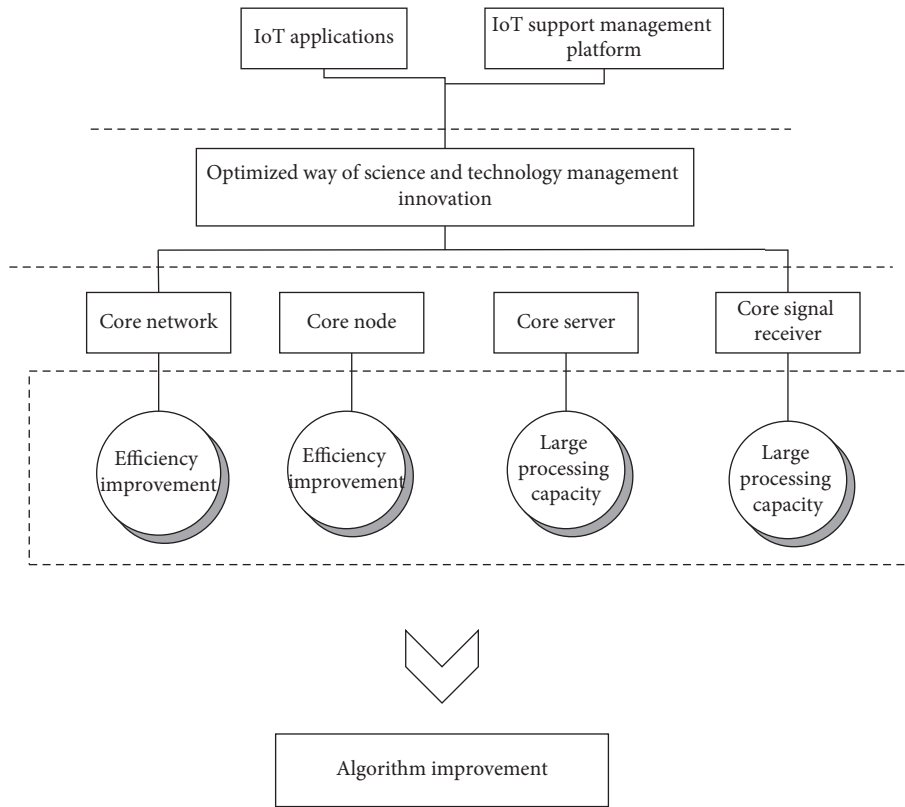
FIGURE 1: General framework of IoT.
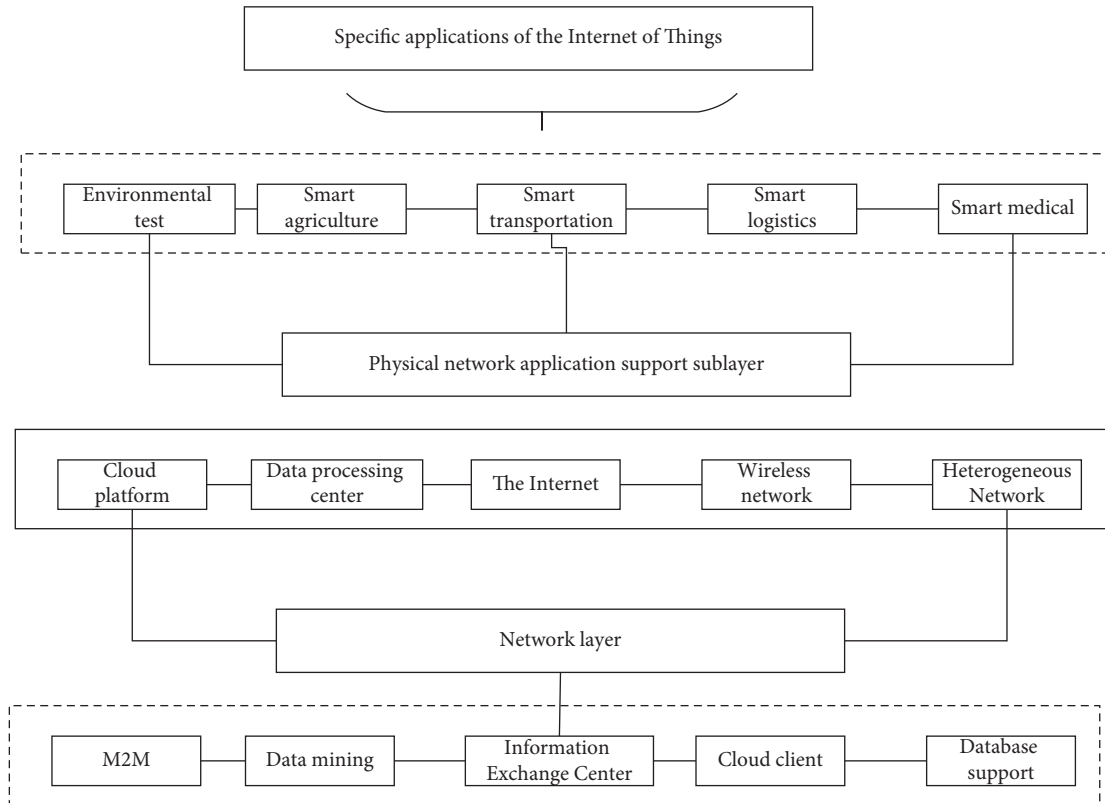
FIGURE 2: IoT technology architecture.
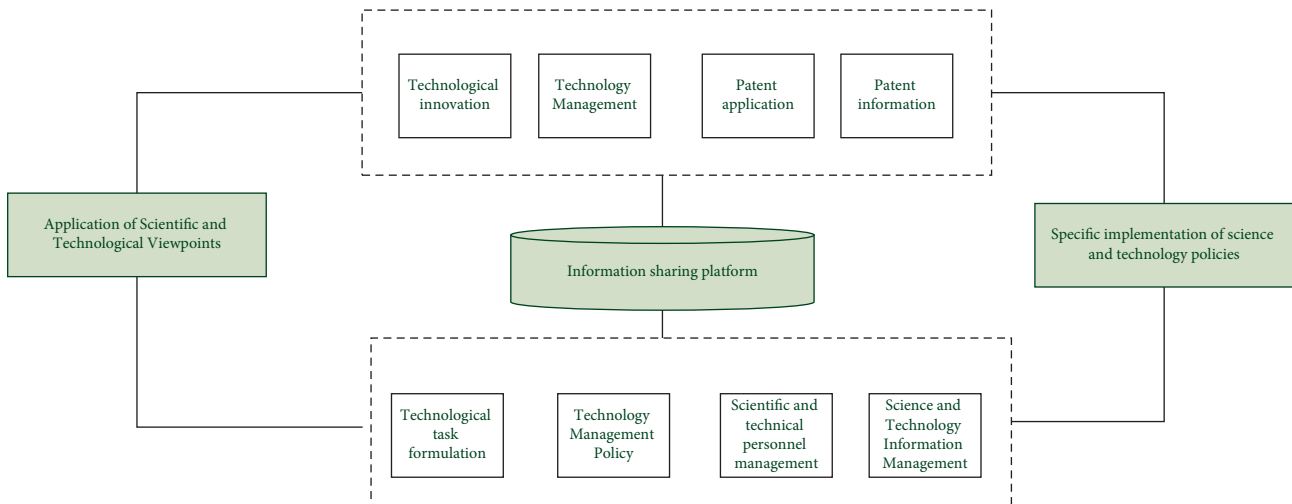
FIGURE 3: Diagram of industrial integration model.

results. The decentralized system, in contrast to the centralized system, usually does not have a macro-high technology management ministry at the national level. This decentralized high technology management can ensure that each government ministry determines the meat of its n work according to its functional areas and management purposes and can respond to various problems promptly, and high technology development can be more W oriented to external needs and maintain high efficiency in the application of high technology results. The dual system is an intermediate type between centralized and decentralized, the national high technology system is generally managed through the national and regional administrative ministries of high technology education, high technology policy is mainly formed through a combination of a top-down and bottom-up approach, and the formation process is accompanied by the participation of many interest subjects, usually the product of the coordination of interests between the systems. With the development of high technology, international competition in high technology has become increasingly fierce, and the demand for innovation in countries has prompted countries to constantly adjust their internal decision-making institutions and organizational frameworks to meet their own internal innovation development needs and adapt to various new requirements. There are also many differences in S&T management systems among countries due to differences in political and economic systems. As one of the many functions of government, high technology management is as important as political, economic, cultural, and social management as a public management function of government. However, due to the special nature of high technology, there are many differences in the ways, systems, and procedures of management compared to other management functions of government.

In summary, the optimal program of high technology management includes the following specific aspects: (1) The planning function. Planning is the prediction of future actions in light of the present situation. High technology planning refers to the actual situation of high technology development as the basis for the future implementation of high technology activities to develop a program. The planning function of government high technology management mainly refers to making plans for future high technology development work by clarifying the objectives of high technology development, reasonably allocating high technology resources, and indicating the direction for the development of high technology, to obtain the maximum promotion of high technology for social and economic development. The planning function of local government high technology management is that the local government, based on implementing and executing the national guidelines, policies, laws, and regulations on high technology development, makes plans, strategies, and policies for high technology development and high technology for economic and social development in a certain period in the future and formulates corresponding laws and regulations according to the actual situation of local high technology development. High technology plans are the fundamental guides to scientific and technological activities carried out by the main body of scientific and technological activities; to more effectively promote the development of local high technology, local governments at all levels have developed a plan for the development of high technology in line with the actual local situation and planning as the basis for scientific and technological activities to improve the level of local scientific and technological development, to better promote the development of the local wild economy. (2) Organizational functions. The organizational function of government high technology management is the premise and basis of its specific high technology management work, but also to achieve the goals of high technology management guarantee. The organization functions of local government high technology management means that the local government through the setting of the appropriate organizational structure, equipped with high technology management personnel, the provisions of high technology management departmental authority, and a clear relationship between the relevant ministries and personnel, achieve specific high

technology management goals. Organizational functions are the preconditions for the realization of other high technology management functions, but also the basis for the orderly conduct of high technology management activities; local governments at all levels to better play the role of high technology management are actively improving its high technology management organizational functions, with the establishment of orderly high technology management operating mechanism to ensure the orderly implementation of government high technology management activities. (3) Coordination functions. The coordination function is an important part of the government's high technology management functions and is the guarantee of the smooth development of high technology activities. The coordination function of local government high technology management includes internal coordination and external coordination; internal coordination is the coordination between government high technology management related ministries, because the government's high technology management functions are scattered in a number of government ministries, in addition to the ministry in charge of high technology management; certain government ministries also have the function of high technology management; in order to ensure the orderly conduct of high technology management, it is necessary to coordinate the relationship between various ministries and agencies. External coordination is to coordinate the relationship between the government and other subjects of high technology activities; there are many subjects of high technology activities; in addition to the government ministries that undertake high technology management functions, universities, research institutions, enterprises, and intermediary organizations that provide services for government high technology management are all subjects of high technology activities. To ensure the smooth development of scientific and technological activities, we must coordinate the relationship between the subjects of scientific and technological activities and clarify the status of each subject in the management of high technology. The government is in a special position in many subjects of high technology activities and has the function of high technology management; therefore, the government must coordinate the relationship between the subjects of high technology activities to provide an orderly social environment for the development of local high technology. After undergoing optimization, its STI efficiency diagram is shown in Figure 4.

4. Experimental Results and Analysis

4.1. Experimental Results. The increase in the number of policy texts involving the departure and on-the-job entrepreneurship of scientific and technological personnel during the period of scientific and technological innovation indicates that the mechanism for the mobility and allocation of scientific and technological personnel in country A is gradually improving, but the proportion of policies on the management of scientific and technological personnel only accounts for 15% compared to other policy categories, indicating that the management of scientific and technological personnel is still weak in the system of scientific and
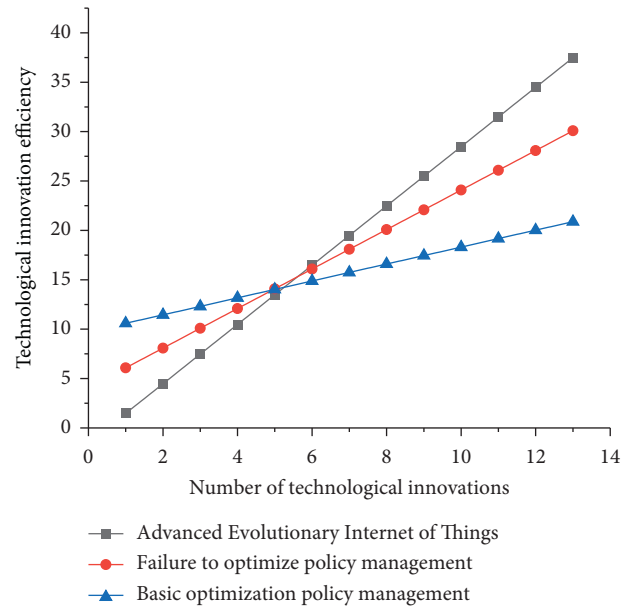
FIGURE 4: Optimized innovation efficiency display chart.

technological personnel in the context of dual innovation, and the mechanism for the selection and use of scientific and technological personnel needs to be improved more. The problems that arise in the management mechanism of scientific and technological personnel are mainly the identity barriers of scientific and technological personnel leaving and starting a business on the job and the ambiguity of the policy of scientific and technological personnel innovation and entrepreneurship. Only by breaking the identity and geographical restrictions of scientific and technological personnel can China's human resources flow and be effectively allocated. Its efficiency diagram after the majorization of high technology management policies is shown in Figure 5.

Subject to the traditional administrative management model, the government's high technology management work has the problem of overcontrolled thinking; the government's high technology management department in the management of high technology will see itself as the leading high technology management, trying to carry out a full range of high technology work, no omission of control. However, the government's high technology management function should be macroscopic guidance, the government's intervention is only to make up for the shortcomings of the market, in high technology activities to give full play to the role of the market to truly promote the development of high technology; the excessive intervention will only make the development of high technology deviate from the normal track, inhibiting the progress of high technology. Therefore, the local government in the management of high technology must establish the concept of service, do a good job in the management of high technology macrocontrol as the main task, and provide a good social atmosphere for the management of high technology, to provide services as the goal and pursuit. After its optimization, there is a significant improvement for high technology innovation, and its quantity changes as shown in Figure 6.
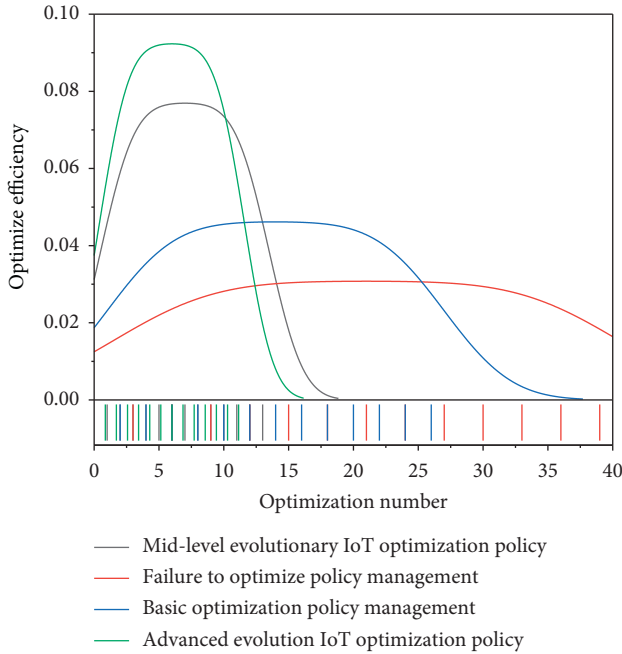
Figure 5: Comparison graph of experimental results.
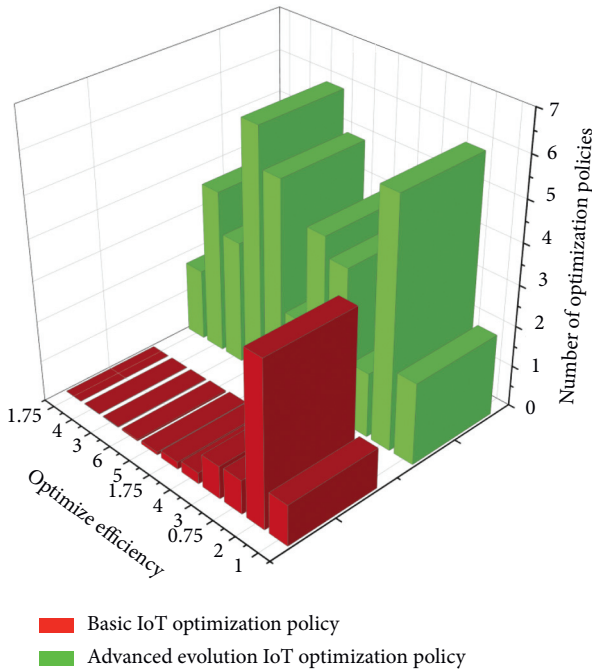
Figure 6: Trends in volume changes.

Figure 7: Trends in the efficiency of the mechanism system.

4.2. Experimental Analysis. To ensure the smooth implementation of high technology management, it is necessary to establish a sharing platform for high technology information, with the ministry in charge of high technology or a specialized agency responsible for summarizing and releasing high technology information on the information platform for sharing by high technology activity subjects. The first step in building a high technology information-sharing platform is to clarify the functional positioning of the high technology information-sharing platform, which should be a comprehensive, vertical, and horizontal exchange information system, and all information and resources related to high technology should be shared on the sending platform.

The government's high technology management department can understand the progress of local high technology work through the information-sharing platform, and other subjects of high technology activities can also understand the government's high technology trends through the information platform, to avoid the blindness of high technology activities and achieve the maximum benefit. Secondly, it is necessary to clarify the main body of the information-sharing platform. The information-sharing platform of high technology is not a way for government ministries to disclose information but should be an information platform jointly constructed by government ministries of high technology management, other relevant ministries, large high technology-based enterprises, universities, and institutes, and relevant enterprises and institutions. The high technology information released by the information platform should be all-rounded, with professional people ± responsible for classifying the high technology information to be released, forming a systematic information network, breaking the regionalization and compartmentalization of information, improving the efficiency of supervision, making up for the information asymmetry problems of ministries and related high technology industries, and eventually forming an information-sharing platform with the local high technology management ministry as the lead and other related high technology activity subjects as the cooperation. The trend in the efficiency of the system of mechanisms formed is shown in Figure 7.

Establishing a pluralistic decision-making advisory system, China's administrative decision-making has long relied mainly on official advisory bodies, and high

first thing is to improve the legal status of high technology assessment by way of legal regulation. One of the reasons why high technology assessment in developed countries can play a good role is that the legal status of high technology assessment is established in the form of laws, and the legal system of high technology assessment is promoted through the formulation of clear legal norms of assessment. Only then can the majorization of the management mechanism of high technology policy innovation driven by the advanced evolutionary IoT calculus reflect its advantages. The algorithm used in the article should pay attention to the efficiency of its operation in the subsequent development process, and the code used by the algorithm should use a more efficient algorithm language to improve the speed and efficiency of the algorithm.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest in this article.

Acknowledgments

References

[1] D. Tomar and P. Tomar, "Dimensionality reduction techniques for IoT based data," *Recent Advances in Computer Science and Communications*, vol. 14, no. 3, pp. 724–735, 2021.

[2] A. Haldorai, S. Murugan, and A. Ramu, "Evolution, challenges, and application of intelligent ICT education: an overview," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 562–571, 2021.

[3] M. Alvi, K. M. Abualnaja, W. Tariq Toor, and M. Saadi, "Performance analysis of access class barring for next generation IoT devices," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 615–627, 2021.

[4] T. Alam, "Cloud-based IoT applications and their roles in smart cities," *Smart Cities*, vol. 4, no. 3, pp. 1196–1219, 2021.

[5] R. Zakaryaei Talouki, N. Javadian, and M. M. Movahedi, "Optimization of the allocation of dynamic vehicle routing with considering traffic," *Journal of Industrial and Systems Engineering*, vol. 13, no. 3, pp. 102–117, 2021.

[6] S. P. A. Datta, T. J. Saleem, M. Barati et al., "Data, analytics and interoperability between systems (IoT) is incongruous with the economics of technology," *Big Data Analytics for Internet of Things*, vol. 2021, pp. 7–88, 2021.

[7] A. Kumar, R. Dhanagopal, M. A. Albreem, and D.-N. Le, "A comprehensive study on the role of advanced technologies in 5G based smart hospital," *Alexandria Engineering Journal*, vol. 60, no. 6, pp. 5527–5536, 2021.

[8] H. Mathur and T. Deepa, "A survey on advanced multiple access techniques for 5G and beyond wireless communications," *Wireless Personal Communications*, vol. 118, no. 2, pp. 1775–1792, 2021.

technology decision-making advice is generally controlled by the Government's Ministry of High Technology Management, and the administrative tendency of decision-making advice is relatively serious. The purpose of establishing an expert consultation system for high technology decision-making is to improve the scientific nature of high technology decision-making, and only by ensuring the independence of the advisory body and establishing a diversified advisory system can the role of expert consultation be truly brought into play. The efficiency trends following the adoption of this method are shown in Figure 8.

5. Conclusion

High technology assessment is professional consultation and evaluation activity carried out by a specialized assessment agency by specific assessment requirements of the client, following certain assessment standards, procedures, and principles, and using scientific assessment methods to assess high technology policies, high technology plans, high technology projects, high technology results, and other relevant high technology activities that need to be assessed. Effective high technology assessment can improve the level of scientific research activities, ensure the quality of scientific and technological achievements, improve the efficiency of the use of scientific and technological resources, and provide a basis for the formulation of relevant high technology policies. A standardized and institutionalized assessment mechanism can greatly improve the quality of high technology assessment, and most developed countries have already built a relatively complete high technology assessment system and made the high technology assessment mechanism an essential part of high technology management and decision-making work.

High technology assessment can effectively guarantee the scientific nature of high technology management, and to give full play to the role of the assessment mechanism, the

[9] J. Mohanty, S. Mishra, S. Patra, B. Pati, and C. R. Panigrahi, "IoT security, challenges, and solutions: a review," *Advances in Intelligent Systems and Computing*, vol. 1199, pp. 493–504, 2021.

[10] L. F. Khalid and S. Y. Ameen, "Secure IoT integration in daily lives: a review," *Journal of Information Technology and Informatics*, vol. 1, no. 1, pp. 6–12, 2021.

[11] J. Ramkumar, M. Baskar, and B. Amutha, "Advanced cognitive models and algorithms," *Cognitive Engineering for Next Generation Computing*, vol. 2021, pp. 121–140, 2021.

[12] L. Tang and Y. Meng, "Data analytics and optimization for smart industry," *Frontiers of Engineering Management*, vol. 8, no. 2, pp. 157–171, 2021.

[13] Q. Cao, "Contradiction between input and output of Chinese scientific research: a multidimensional analysis," *Scientometrics*, vol. 123, no. 1, pp. 451–485, 2020.

[14] L. Ganushchak-Efimenko, I. Hnatenko, R. Kozhushko, R. Rębilas, V. Rubezhanska, and T. Krakhmalova, "Optimisation models of investment management in the activities of innovative enterprises," *Management Theory and Studies for Rural Business and Infrastructure Development*, vol. 42, no. 3, pp. 225–234, 2020.

[15] M. A. Sadeeq and S. Zeebaree, "Energy management for Internet of things via distributed systems," *Journal of Applied Science and Technology Trends*, vol. 2, no. 2, pp. 59–71, 2021.

[16] A. Affandi, A. S. Sarwani, H. Erlangga, and A. Purwanto, "Optimization of MSMEs empowerment in facing competition in the global market during the COVID-19 pandemic time," *Systematic Reviews in Pharmacy*, vol. 11, no. 11, pp. 1506–1515, 2020.

[17] K. Li, T. Zhou, and B.-h. Liu, "Internet-based intelligent and sustainable manufacturing: developments and challenges," *International Journal of Advanced Manufacturing Technology*, vol. 108, no. 5-6, pp. 1767–1791, 2020.

[18] T. E. T. Dantas, E. D. De-Souza, I. R. Destro, G. Hammes, C. M. T. Rodriguez, and S. R. Soares, "How the combination of circular economy and industry 4.0 can contribute towards achieving the sustainable development goals," *Sustainable Production and Consumption*, vol. 26, pp. 213–227, 2021.

[19] J. R. Saura, "Using data sciences in digital marketing: framework, methods, and performance metrics," *Journal of Innovation and Knowledge*, vol. 6, no. 2, pp. 92–102, 2021.

[20] K. H. Zhang, "Industrial policy and technology innovation under the US trade war against China," *The Chinese Economy*, vol. 53, no. 5, pp. 363–373, 2020.

Research Article

Optimal Analysis of Human Resources Allocation in Colleges and Universities Based on Internet of Things Technology

Ai Zhang [iD],[1,2] Zhengmao Ju,[3] and Yiling Liu[2]

[1]*Gannan University of Science and Technology, Jiangxi, Ganzhou 341000, China*
[2]*College of Applied Sciences, Jiangxi University of Technology, Jiangxi, Ganzhou 341000, China*
[3]*Mokwon University, Daejeon 34705, Republic of Korea*

Correspondence should be addressed to Ai Zhang; 9320070268@jxust.edu.cn

This article takes the development and superiority of the human resource system under the IoT technology as the background and adopts the popular IoT architecture and the current latest technical theories to develop and improve the human resource management system. First, we introduce the establishment of the J2EE development environment and then introduce the realization of the three layers of the overall IoT architecture of the system, namely, the realization of the presentation layer, the realization of the layer, and the realization of the data and then the realization of the main modules of the system; finally, we explain the function test of some modules of the system. After that, we mainly analyze the requirements of the system, first the overall requirements analysis and then the function and nonfunctional requirements analysis of the system. On the basis of perfecting the service-oriented modeling of manufacturing resources, the integration and optimization of resources are carried out according to the requirements of manufacturing tasks, and the discovery mechanism and matching methods of manufacturing resource services are studied to reduce the solution space of resource allocation. Aiming at the problem of optimal utilization of manufacturing resources, this paper establishes a scientific and reasonable evaluation index system to reduce resource service costs and improve resource utilization. At the same time, after analyzing the problem of real-time distribution of resources in the manufacturing Internet of Things environment, the real-time information of resource distribution is designed. The interactive mechanism and a two-tier optimization method based on real-time information-driven resource distribution tasks are proposed. The simulation experiment results show that the optimization method proposed in this paper, compared with the traditional resource distribution method, not only reduces the carrying distance of human resource distribution but also reduces the empty load rate, thereby reducing the waste of human resources.

1. Introduction

With the help of this technology, traditional human resource management is also facing unprecedented opportunities and challenges [1]. In enterprises, human resource management has become an indispensable part of enterprise management. Manual human resource management is time-consuming and laborious, prone to errors, cumbersome, and inefficient, and the existing personnel system is highly coupled with the enterprise [2]. Therefore, the previous management model has changed and been replaced by the information network management model. With the penetration of the concept of the Internet of Things in the manufacturing field, the development process of manufacturing companies has evolved from the traditional "black box" model to a "three-dimensional space plus time multidimensional, transparent ubiquitous perception" model, which is driven by information perception technology [3–5]. This article uses advanced computer technology to develop enterprise human resource management system, networked and informatized human resource management. Such a system can reduce the work burden of enterprise employees, move employees' positions in real time and efficiently according to local conditions, maximize employees' potential, and improve the economic benefits of the enterprise [6–8].

Li et al. [9] proposed a shift in the focus of corporate personnel information management, which marked the emergence of research on the theory and practical application of human resource management. The human resource management system developed by Oracle mainly reduces the cost of enterprise management through system analysis, management mode, and establishment of employee incentive mechanisms. The human resource management system developed by IBM mainly provides solutions and management services. The recruitment of teachers in space is open, and the teacher data and other employment information of the schools are very transparent, which has achieved a reasonable flow of talents [10–13]. Chen et al. [14] proposed a new service-oriented model of physical entities and proposed a method of secondary combination reasoning to speed up the combination process. In the same year, they further proposed a new context-aware resource-led service model. Awoyemi et al. [15] studied the extended OWL-S model to describe the contextual characteristics of service providers in CPS and also studied technology for generating task demand planning. Fang et al. [16] conducted research on dynamic material distribution in the workshop under the IoT environment. Aiming at the problems of low transparency in the production process, difficulty in collecting material information, low material distribution efficiency, and poor information interaction, they proposed a real-time sensing information-driven. Fischer et al. [17] studied the application of the Internet of Things technology in the disaster relief material distribution management system, using the real-time monitoring data collected by the Internet of Things to analyze and judge the results and play a comprehensive auxiliary decision-making process for the entire material distribution and transportation process. Some scholars studied the material distribution and monitoring system of mixed-flow assembly workshops and put forward the overall framework of material distribution and tracking in response to the requirements of material distribution and monitoring on the production site and assembly error prevention [18]. Since the status information of each distribution carrier, pallet, and so forth (such as location, material category, and quantity) can be sensed in real time, the description of dynamic resources based on real-time process perception can be further optimized. In addition, by constructing a real-time information interaction mechanism and perfecting the optimization of real-time information-driven resource distribution tasks, it is also a problem to be finished in resource distribution in the manufacturing Internet of Things environment [19–21].

In order to cope with the future expansion of the functions of the Internet of Things enterprises, this system adopts J2EE technology, B/S platform mode, and SSH framework to develop the human resource deployment system. Logging in to office is in the concept of cloud office. At the same time, the ULHA evaluation model is introduced into the enterprise human resource management system to better help the enterprise evaluate the company's employees reasonably, conduct comprehensive and humanized management, and help companies make better decisions in the recruitment process. Due to the complexity of the regression model, it cannot be done by relying solely on the spatial analysis function of GIS. The regression analysis must be done with the help of professional statistical analysis software SPSS. Firstly, the background and significance of the system research and the main existing problems are explained, and the related technologies and theories of the system are introduced; secondly, the functional requirements and nonfunctional requirements of the system are analyzed; and finally, the main modules of the system are implemented and tested, and the test results basically meet the design requirements of the system. The developed system is easy to operate, simple to use, and easy to expand, which improves the operation efficiency and market competitiveness of the enterprise and at the same time improves the enthusiasm and work efficiency of employees.

2. Optimization Analysis Model of Human Resource Allocation in Colleges and Universities Based on Internet of Things Technology

2.1. Hierarchical Structure of IoT Technology. At the technology level of the Internet of Things, the B/S model is a relatively popular software architecture. It is also a major trend in developing the Internet of Things architecture for large-scale systems in the future. In the B/S mode, users only need to install a browser to directly access the server and exchange data [22–25]. The three-tier system structure of B/S includes a presentation layer, a logical business layer, and a data layer. Figure 1 is the hierarchical topology of the IoT technology.

In the B/S IoT architecture, this article mainly adopts an open-source Internet of Things architecture that is relatively mainstream, relatively advanced in technology, and relatively mature and stable developed products. In other words, the SSH Internet of Things architecture is used to design and develop. According to the different needs of the system, this paper encapsulates the different subsystems, respectively. The advantage of this is that it is conducive to the integration of the various subsystems after the system development is completed:

$$f_{11} = \left(\frac{1}{n} \times \sum_{i=1}^{n} x_i^2, \frac{1}{n} \times \sum_{i=1}^{n} y_i^2 \right),$$

$$c = \Pr(c \mid o) \times \Pr(o) \times \text{IoU} = \Pr(c) \times \text{Iou}. \tag{1}$$

The technology Internet of Things architecture adopted by J2EE mainly adopts the idea of low-coupling hierarchical design and the distributed system Internet of Things architecture, encapsulating different logical structures into different application services for processing, and fully embodies the hierarchical design idea of the system. The presentation layer adopts the framework for design, the business layer mainly adopts the framework, and the data persistence layer adopts the Hibernate framework, which guarantees the system's stability as a whole. In the system development, the Spring structure, Struts structure, and
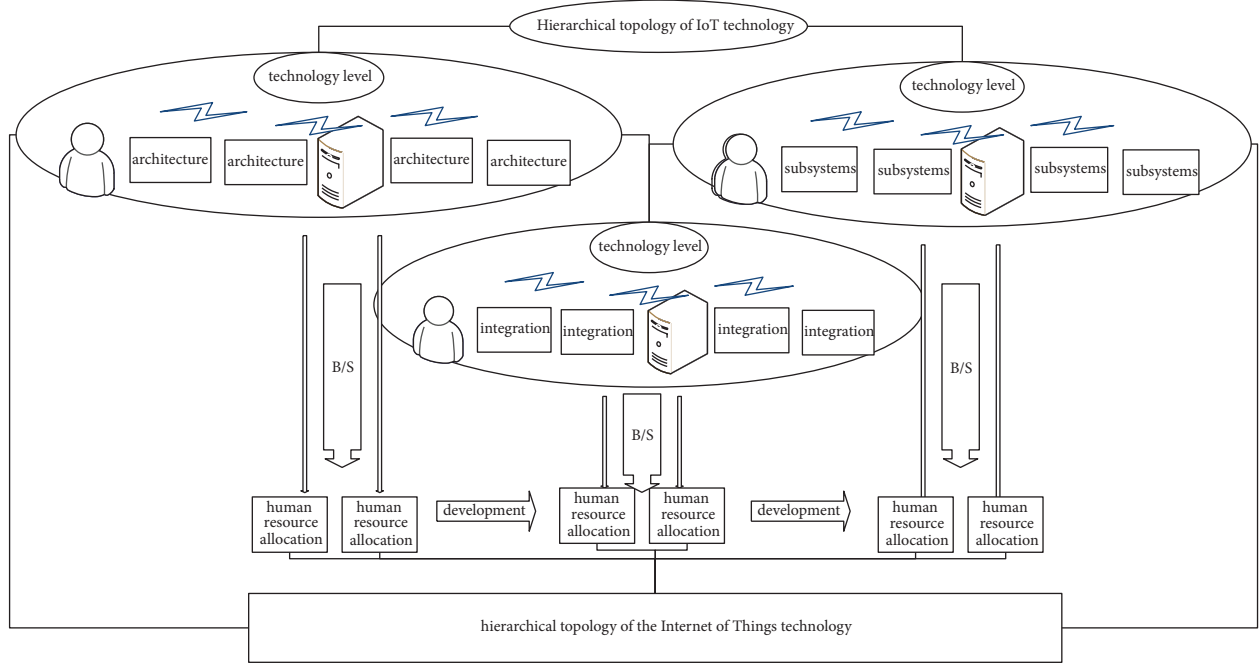
Figure 1: Hierarchical topology of IoT technology.

Hibernate framework have become the main flow frameworks for developing a lightweight framework system:

$$l = \alpha \times \sum_{i=0}^{S^2} \sum_{j=0}^{B} \left[(x_i - \overline{x})^2 + (y_i - \overline{y})^2 + (x_j - \overline{x})^2 + (y_j - \overline{y})^2 \right],$$

$$\psi(x) = \frac{f(x,d) - f(x,m)}{f(x,d) + f(m,d)}.$$

$$(2)$$

It can serve the functions of heterogeneous resources, provide services to the outside world with a unified and universal interface, and connect to the manufacturing platform in a loosely coupled manner to realize the full interconnection of the manufacturing services of physical manufacturing resources. However, the current manufacturing resource model mainly models the processing capability of the resource and less considers the real-time status and location of the resource, making it difficult to dynamically optimize based on the real-time status of the manufacturing resource:

$$\chi(x) = \begin{cases} e^{-\lambda L \times (1 + \lambda x^2)}, & \chi \geq \chi(x), \\ 0, & \chi < \chi(x), \end{cases}$$

$$G(x) = \begin{cases} g(x) \times [0, x, 1], & x \subset X, \\ g(x) \times [x, 1, 0], & x \subset Y, \\ g(x) \times [1, 0, x], & x \subset Z. \end{cases}$$

$$(3)$$

The real-time manufacturing information acquisition method is the premise and foundation for realizing the service-oriented modeling of dynamic manufacturing resources. According to its information structure, it is hierarchical, structurally related, and gradually superimposed and merged. It can form the basic information of the production process that has practical meaning for manufacturing resources. Then, the information is sorted and combined to form closely related production tasks and production processes. Real-time multisource manufacturing data provide data support for the production information system.

2.2. Human Resource Allocation Algorithm.

Such an Internet of Things architecture has relatively unified specifications and standards on system components, technical levels, and various service Internet of Things architectures. The advantage of this is that, for different platforms that also follow the J2EE Internet of Things architecture, that is, between heterogeneous platforms, the system performs well in terms of compatibility, allowing the enterprise to communicate well internally and externally. After SPSS software completes logistic regression analysis, the output digital matrix can also be converted into raster graphics in GIS software, thus solving the problem of mutual import of data in GIS and professional statistical analysis software. In the framework of J2EE technology, programmers only need to develop large-scale enterprise-level application systems according to the corresponding specifications, without considering the underlying hardware issues. Considering the compatibility issues between different J2EE platforms and products, different J2EE providers will also support the corresponding standards drawn up in different J2EE versions. In order to facilitate the definition and the expression of operators, this

article gives a zero-centered, odd-numbered language term set Hs = {$h3$ = excellent, $h2$ = very good, $h1$ = good, $h0$ = fair, $h-1$ = bad, $h-2$ = very bad, $h-3$ = very bad}. In order to avoid the loss of decision information in language term calculations, language terms are generally expanded. For the language term set of this article, this article adopts the following expansion scale, where q is a sufficiently large natural number.

Table 1 is the description of human resource allocation. This layer communicates directly with the database storing the data and can complete the functions of inserting, deleting, modifying, and querying data. When the business logic layer issues a request to the database, the layer responds accordingly. The function of this layer is mainly composed of two functions of OR mapping and database operation layer. It adopts the Hibernate Internet of Things architecture and is based on the OR Mapping framework to carry out persistent package management on the data operation of the system and realize the data and business data of the database. The use of such a framework does not need to consider which database is used, which can effectively ensure the independence of the system. The most resource-consuming system is access to the database operation. Therefore, to improve the content and response performance of the data, the operation of the data access layer on the database must be optimized. The application and realization of the ontology concept in the Web require a standard, common, and standardized language that can be used by computers to describe services. It uses ontology concepts to express the semantic information of Web services and accurately describes the meaning of concepts in a way that can be understood by computers to realize service discovery.

2.3. Model Optimization Analysis.

Separating the data model of the Internet of Things from the code of the user interface is the main purpose of applying the MVC pattern. According to the MVC Internet of Things architecture model, the application is separated from input, application processing, and output through high-level interfaces between objects to ensure the decoupling between the various parts of the application system and, at the same time, make each layer of the system relatively independent of processing its own business. Its main components include BeanTags, HTMLTags, Logic Tags, NestedTags, and Template Tags. The model is mainly responsible for identifying the state of the system. In this case, the method of using a stepwise linear statistical model will cause large errors in the results and even unreasonable results. As far as the structure of the entire system is concerned, ActionFormBean is used to represent a nonpersistent state. For some persistent states, the Utitle package is used, which is included in Struts itself. The advantage of this is that it becomes more convenient when operating with the database. In the Struts framework, ActionServlet plays the role of the controller, but it is not completely involved in business logic and requires interaction between Action, ActionMapping, and ActionForward. The roles and tasks of the three components are different. The main responsibility for the realization of

Table 1: Description of human resource allocation.

Index	Dataflow rate (%)	Expansion scale	Factor weight
1	17	1.21	0.38
2	14	2.01	0.09
3	25	1.89	0.12
4	31	1.77	0.41

business logic is action. The difference of business logic is represented by ActionMapping, which is to distinguish. The direction of the process is represented by ActionForward.

Figure 2 is the logical framework of model optimization analysis. In the manufacturing process, when the sensor node perceives the corresponding event (data), the designed event-driven message transmission mechanism firstly describes the event information collected by the sensor node on the resource side as having {ID, Event, Data, Time} a collection of data structures. Then, the mapping relationship F between different types of sensors and different types of manufacturing information is established in the background. Each sensor event is monitored through multithreading technology, combined with the event-driven IoT architecture (EDA) and mapping table to encapsulate and transmit the data captured by the event on the sensor. The static attributes of the manufacturing resource service give a unique identifier to the service, which summarizes the service capabilities. When a manufacturing task requires a query service, if the query content is a keyword, you can directly query through the UDDI registration center to obtain the universally unique identifier (UUID) of the manufacturing resource service that meets the requirements and return it to the information system. The matching engine finds similar concepts in the ontology database according to the semantic description, calculates the matching degree, obtains the manufacturing resource service that meets the requirements, and finds the UUID of the service in the UDDI registry and returns it to the information system.

2.4. Data Weighting Factor Recursion.

The human resource deployment data Web is an extension of traditional Web services. The resource information in the Semantic Web is no longer a simple information location, but a real way to understand the meaning of information, improve the search, and get high efficiency and intelligence of computers. For the meaning of information on Web computers, the relevant domain knowledge and relationships represented by the information need to be clearly defined. Not only can ontology provide clear definitions of related domain knowledge and mutual relations, but also the logical reasoning capabilities it supports are more conducive to computer understanding and processing. Therefore, the ontology concept model has become the foundation and core of the development of Semantic Web technology. The performance appraisal data comes from the performance appraisal scoring table and relevant system standard benchmarking and original work records. The level of work completion is determined according to the completion of the performance
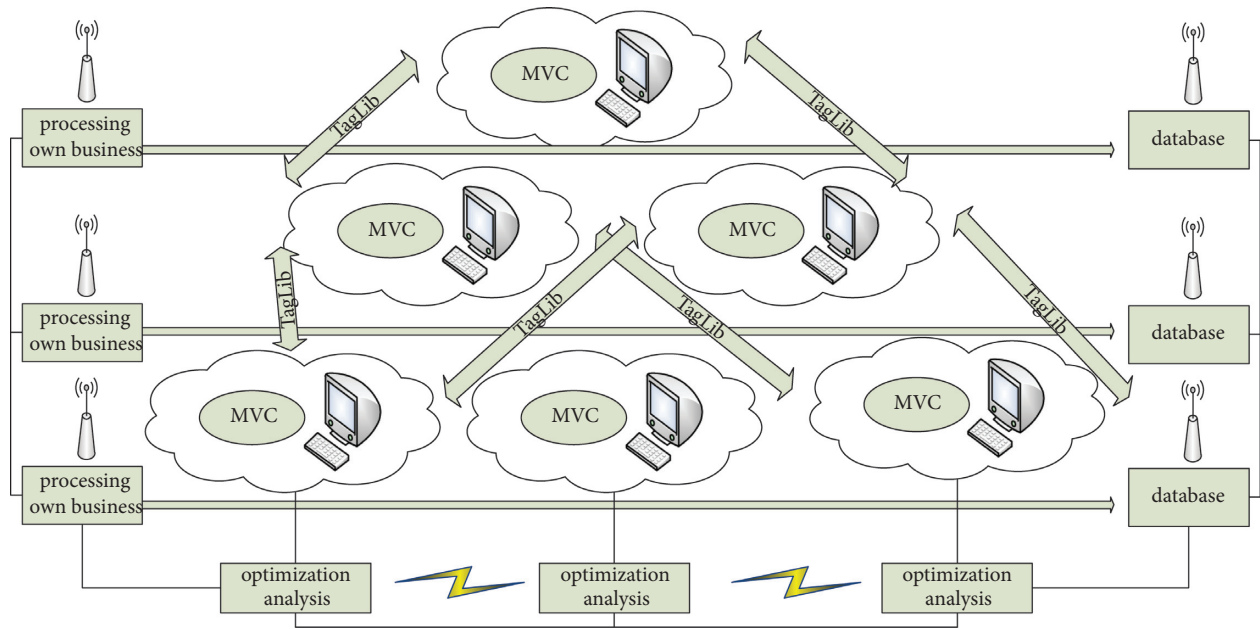
FIGURE 2: Logic framework of model optimization analysis.

appraisal indicators and is divided into four levels. The actual completion of the work becomes the actual completion of the responsibility objectives and work tasks during the assessment period. The actual assessment scores are scored by the assessment unit (assessor) based on the employee's responsibility target and the completion value of the work task. Figure 3 shows the classification of data weighting factors.

Through the information collected by the Internet of Things, this article should take an employee by affecting task assignment. When these elements are available, it is not necessary to assign tasks only based on distance. Reasonably, in addition, the Internet of Things will collect data for the frequency distribution, and processing time distribution of emergency resource allocation events in each area can be obtained. We form an effective and comprehensive database, analyze and summarize the data in the database, and slowly find the human resources law from the data, which can provide a scientific basis for future human resources work.

3. Results and Analysis

3.1. Analysis of IoT Data Characteristics. In order to reflect the status quo of H company's human resource value chain, this paper uses Schuster's human resource index method to conduct a questionnaire survey on H company. The scale used in this method is designed with 64 questions and 15 dimensions. The dimensions, respectively, reflect the main and status quo of the four links of the human resource value chain management model, the relationship between the 15 dimensions of the human resource index, and the 4 dimensions of the human resource value chain, so that the human resource index survey can reflect the management status of the company's human resource value chain. The results show that the proportion of disasters in the 1 and 2 divisions
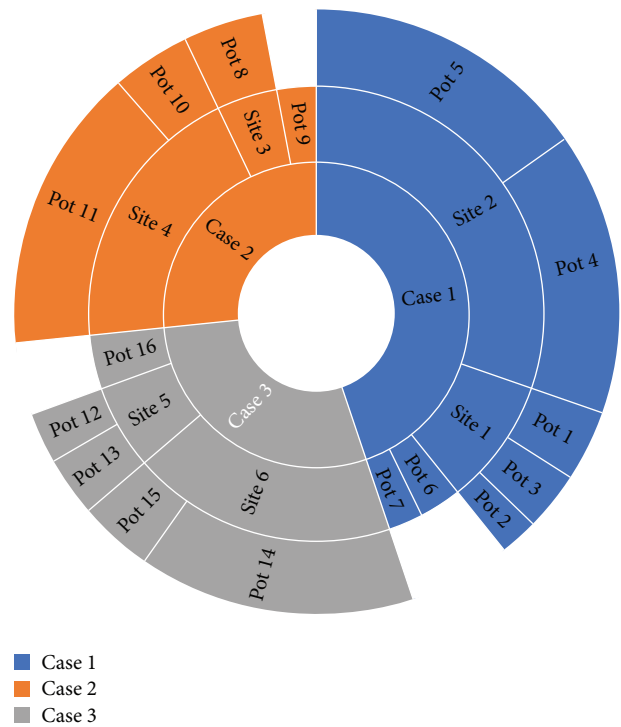
FIGURE 3: Classification of data weighting factors.

representing low levels of risk has decreased significantly, while the proportion of disasters in the 4 and 5 divisions representing high levels of risk has decreased. The questionnaire adopts the Schechter 5-point scale model, which enables the respondent to answer quickly and conveniently. A total of 112 questionnaires were issued to H company, 105 valid questionnaires were received, and the effective recovery rate was 93.3%. There are 64

questions in the questionnaire, divided into 15 dimensions, and scored using a 5-point scale. The reliability coefficient (Alpha) of this scale is between 0.757 and 0.929.

Figure 4 is the statistical distribution of the reliability coefficient of the Internet of Things optimization. It can be seen that the scores of each latitude of the H company's human resource index are generally low. Among them, the highest score is the organizational environment 3.45, indicating that the organizational environment of the employees of H company is relatively harmonious. The second highest score is organizational goal 3.25 and group collaboration, indicating that the employees of H company have done a relatively good job in improving information communication, group collaboration, and so on. The lowest score is the remuneration system, with a score of 2.58, indicating that H company's human resources compensation management is lacking, which leads to low employee satisfaction; the second-lowest score is the management quality and the overall management quality of H company is not high. In terms of recruiting and training new recruits, H company lacks reasonable measures and talent mining and value creation. The company does not conduct an in-depth analysis of the company's positions and lacks direction in the company's strategic planning. At the same time, it also reflects that the value evaluation link in the H company's human resource value chain management is not well done, and a new understanding and improvement of the overall management of the organization is needed.

3.2. Simulation of Human Resource Allocation.

The company expanded related business and elected the new executive director from the four main persons in charge. This article assumes that these four persons are A, B, C, and D. The current work mainly evaluates these four persons. The leaders, colleagues, and subordinates of these four people are evaluated in language terms, and the leaders, colleagues, and subordinates are used as representative groups to participate in the evaluation. The evaluation value of each group selects the group's evaluation range for the person in charge. Based on these theories, the four main persons in charge from the three levels finally obtained the following evaluation information based on the uncertainty language evaluation scale. The city is divided into several large districts, and each district is divided into several blocks. Each block is assigned information collectors and personnel. Therefore, the simulation system is simulated according to the district, and the simulation system is mainly divided into three modules. The first part is the data initialization module. Before the simulation experiment, you need to input the corresponding parameters. The second part is the simulation experiment algorithm part, which is divided into a scheduling module and a configuration module, used to solve the scheduling and configuration problems in the current process and ensure that the simulation experiment normally runs according to the designed logic. The third part is the data interaction module, which interacts with the database, writes necessary data, and reads the required information for queries. Table 2 is the data interaction module of human resource deployment.
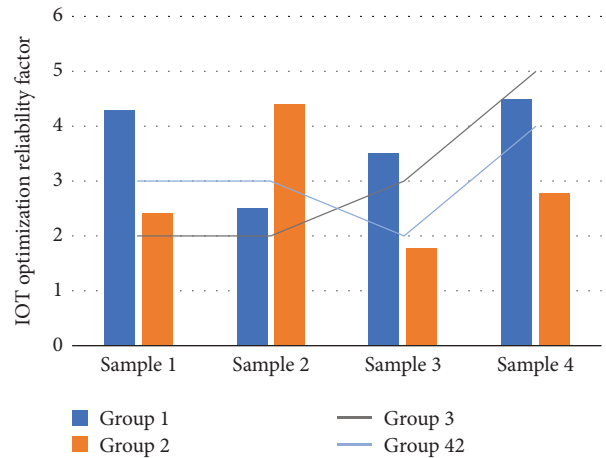
Figure 4: The statistical distribution of the reliability coefficient of the Internet of Things optimization.

Using the ULHA operator in the evaluation of human resources, compared with not using the ULHA operator, the evaluation is more accurate and scientific. It has been used in many human resource management systems and has achieved very good results. The calculation accuracy is more than 10% higher than the traditional method, and this mode is simple to operate, suitable for general scenarios. First of all, the overall index M and the change trend of the employee load rate are the same as the previous part. We use the performance appraisal system of the cloud platform to improve the performance management efficiency of the enterprise; the value evaluation process is to check the implementation results of the above two links in the value chain. Salary management is particularly important. Salary management under the cloud platform makes employees more challenging and fair. The final value distribution process has an incentive effect on the entire value chain and is the basis for retaining valuable employees, as shown in Figure 5.

Through the establishment of H company's exclusive HR-cloud system, the cloud system is deployed on the Internet and positioned as a full-staff service, giving full play to the real value of the human resources value chain system and realizing the line of core modules such as personnel, salary, attendance, and performance. At the same time, it can realize employee self-service business processing, full participation in performance management, self-service reporting of training needs, and so on so that the human resources business can truly provide services for senior executives, line managers, and employees. We integrate the human resources of H company and optimize the allocation of company's H human capital as a whole. Through the HR-cloud human resources value chain management platform, employees can timely understand the human resources policies and regulations formulated by the upper leaders of H company; it can supervise the daily human resource management business of various departments to ensure the management system, which greatly improves the satisfaction of employees.

TABLE 2: Data interaction module for human resource deployment.

Number	Description	Value
1	Type, resource, quality, volume	20
2	Model, load capacity, module volume, transportation distance	40
3	Distribution demand, distribution frequency, distribution point	20
4	Module resource distance, number of modules used, delivery time limit	40
5	Meet time resource requirements, meet customer distribution needs	50
6	Distribution center to customer path, customer to customer path	70
7	Gathering goods, completion of module distribution tasks	60
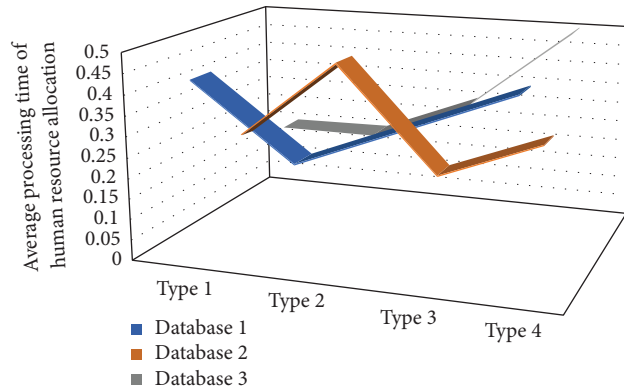
FIGURE 5: Average processing time of human resource allocation.

3.3. Analysis of Experimental Results. The human resources performance appraisal indicators are uniformly set as follows: the numbers of performance appraisal indicators for the three types of management personnel of middle management, supervisor, and management assistant are set to 10, respectively; the number of performance appraisal indicators for technical personnel is set to 12; and the numbers of performance appraisal indicators for senior level, standard level, and junior/other three types of production and operation service personnel are set to 10, respectively. The performance evaluation indicators of various types of personnel are divided into three categories: work performance, work behavior, and workability. Each indicator clarifies the specific evaluation content and items. Randomly select the local area model unit, and use the CF value of each factor data type as the independent variable of the logistic regression model and the risk as the dependent variable. Import the data into the SPSS statistical analysis software, and use binary logistic regression analysis. The evaluation weights and standard scores are determined according to the job analysis and job evaluation results, and the evaluation is specified. In the future, according to the company's overall development requirements, performance management needs, and employee development needs, the employee performance appraisal indicator system will be adjusted in a timely manner. The hierarchical overall analysis is used to calculate the weight of the bottom element to the target layer element, that is, the weight of the delivery distance, the priority of the delivery task, and the utilization of the delivery resource for the optimal task plan. The largest eigenvalue of the judgment matrix G1 is an important parameter for consistency testing. The largest eigenvalue corresponding to the eigenvector (orthogonal normalization vector) is called the weight vector. Each component corresponds to the delivery time, the priority of the delivery task, and the resource. Figure 6 is a histogram of human resource performance appraisal indicators.

Based on the data accumulated by the development of network services, through data processing and analysis, real-time dynamic monitoring of operating conditions collect the dynamic change data of the online public human resources market, implement analysis, display the operation of the online public human resources market, and provide dynamic supervision functions. Due to the lack of opportunities for employees of H company to participate in management, the organizational structure of the Internet industry tends to be flat, with low management personnel, and employees rarely participate in management. However, to realize the potential of employees, it is possible to appropriately increase the opportunities for employees to participate in the management and explore technical skills. In this regard, H company urgently needs improvement in this regard.

Figure 7 shows the distribution of real-time dynamic supervision and operation of the Internet of Things. In order to verify the research on the optimization of dynamic resource distribution tasks in the manufacturing Internet of Things environment, this section designs the following manufacturing workshop material distribution scenarios. The proportion of excellent and competent employees reaches 53%. These personnel have good job competence and promotion potential and can be used as future talent reserves. Incompetent personnel belong to the department of the R&D center. The out-buffer is used to store materials
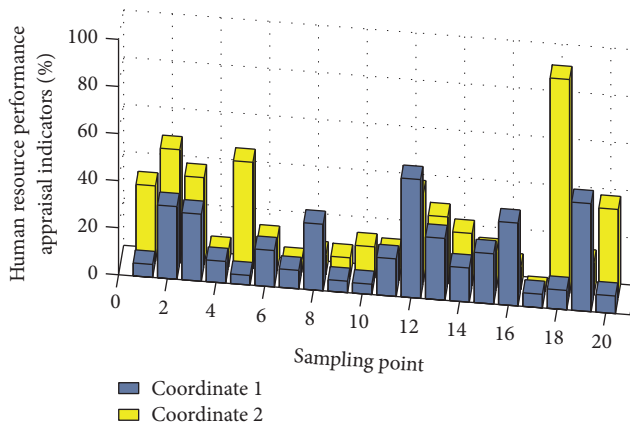
Figure 6: Histogram of human resource performance appraisal indicators.
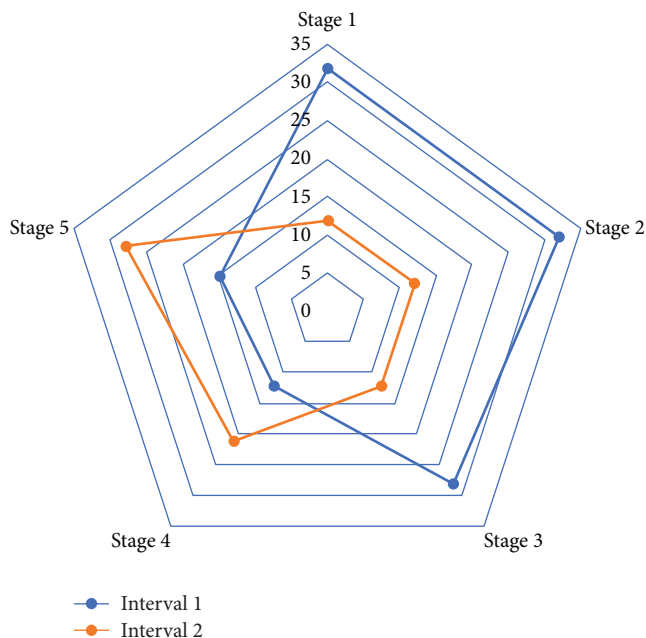
Figure 7: Distribution of real-time dynamic supervision and operation of the Internet of Things.

that have been processed by the workstation and used in other manufacturing links; the second is the storage area, which contains a series of three-dimensional shelves to store and maintain the normal workshop; and the third is a forklift, which has real-time perception service capabilities by configuring sensors and RFID and other equipment. At the same time, relevant hardware devices are also attached to key locations and resource storage areas to collect real-time distribution information, such as attaching RFID tags at key intersections, material buffers, and shelves in the workshop to collect location information.

4. Conclusion

This system mainly uses the B/S Internet of Things architecture for design and development. First, it can effectively

protect the security of sensitive data, especially when adding and modifying records to the database. It can make general functions easier to use and ensure that complex functions interact. At the same time, it focuses on the problem of service modeling based on real-time information, service combination optimization, and resource distribution optimization of dynamic manufacturing resources in the manufacturing Internet of Things environment. Under the guidance of the resource-as-a-service ideology, a manufacturing resource modeling system with real-time perception service capabilities is organized, and the manufacturing resource contextual information modeling method, manufacturing resource servicing modeling method, and manufacturing resource service combination tasks are detailed respectively. This module uses a histogram to analyze the comparison between the existing emergency source capacity of the district and the optimized configuration. The modeling method establishes the research foundation for the dynamic manufacturing resource service combination optimization and resource distribution optimization later. In the manufacturing Internet of Things workshop, human resources, material resources, and equipment resources are the main research objects, a context-aware manufacturing resource service discovery mechanism is designed, a service composition model based on workflow diagrams is established, and gray correlation is adopted. The proposed human resource allocation plan simulation system uses the historical data collected in the smart city to deal with emergency human resources problems in various regions for simulation experiments, which can help city managers find the plan for each region based on their own needs to deal with emergency human resources issues.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this article.

Acknowledgments

References

[1] M. Hawkins, "Cyber-physical production networks, internet of things-enabled sustainability, and smart factory performance in industry 4.0-based manufacturing systems," *Economics, Management, and Financial Markets*, vol. 16, no. 2, pp. 73–83, 2021.

[2] G. Yan, "Simulation analysis of key technology optimization of 5G mobile communication network based on internet of things technology," *International Journal of Distributed Sensor Networks*, vol. 15, no. 6, p. 155, 2019.

[3] D. Jiang, "The construction of smart city information system based on the internet of things and cloud computing," *Computer Communications*, vol. 150, pp. 158–166, 2020.

[4] Y. Ding, M. Jin, S. Li, and D. Feng, "Smart logistics based on the internet of things technology: an overview," *International Journal of Logistics Research and Applications*, vol. 24, no. 4, pp. 323–345, 2021.

[5] F. J. Ferrández-Pastor, J. M. García-Chamizo, M. Nieto-Hidalgo, and J. Mora-Martínez, "Precision agriculture design method using a distributed computing architecture on internet of things context," *Sensors*, vol. 18, no. 6, Article ID 1731, 2018.

[6] H. Sándor and G. Sebestyén-Pál, "Optimal security design in the internet of things," in *Proceedings of the International Symposium on Digital Forensic and Security*, pp. 4–6, Tirgu Mures, Romania, April 2017.

[7] P. Liu, W. Qingqing, and W. Liu, "Enterprise human resource management platform based on FPGA and data mining," *Microprocessors and Microsystems*, vol. 80, p. 103, 2021.

[8] R. Singh, A. Gehlot, S. V. Akram et al., "Cloud manufacturing, internet of things-assisted manufacturing and 3D printing technology: reliable tools for sustainable construction," *Sustainability*, vol. 13, no. 13, Article ID 7327, 2021.

[9] Y. Li, A. Alqahtani, E. Solaiman et al., "IoT-CANE: a unified knowledge management system for data-centric Internet of Things application systems," *Journal of Parallel and Distributed Computing*, vol. 131, pp. 161–172, 2019.

[10] T. Qu, Y. Pan, X. Liu et al., "Internet of things-based real-time production logistics synchronization mechanism and method toward customer order dynamics," *Transactions of the Institute of Measurement and Control*, vol. 39, no. 4, pp. 429–445, 2017.

[11] M. Agiwal, N. Saxena, and A. Roy, "Towards connected living: 5G enabled internet of things (IoT)," *IETE Technical Review*, vol. 36, no. 2, pp. 190–202, 2019.

[12] M. A. M. Sadeeq and S. Zeebaree, "Energy management for internet of things via distributed systems," *Journal of Applied Science and Technology Trends*, vol. 2, no. 2, pp. 59–71, 2021.

[13] I. Alam, K. Sharif, F. Li et al., "A survey of network virtualization techniques for internet of things using SDN and NFV," *ACM Computing Surveys*, vol. 53, no. 2, pp. 33–40, 2020.

[14] L. Chen, P. Zhou, L. Gao, and J. Xu, "Adaptive fog configuration for the industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 4656–4664, 2018.

[15] B. S. Awoyemi, A. S. Alfa, and B. T. J. Maharaj, "Resource optimisation in 5G and internet-of-things networking," *Wireless Personal Communications*, vol. 111, no. 4, pp. 2671–2702, 2020.

[16] Y. Fang, Q. Chen, N. N. Xiong, D. Zhao, and J. Wang, "RGCA: a reliable GPU cluster architecture for large-scale internet of things computing based on effective performance-energy optimization," *Sensors*, vol. 17, no. 8, Article ID 1799, 2017.

[17] G. S. Fischer, R. da Rosa Righi, G. de Oliveira Ramos, and C. André da Costa, "ElHealth: using internet of things and data prediction for elastic management of human resources in smart hospitals," *Engineering Applications of Artificial Intelligence*, vol. 87, Article ID 1032, 2020.

[18] P. Gibson, "Internet of things sensing infrastructures and urban big data analytics in smart sustainable city governance and management," *Geopolitics, History, and International Relations*, vol. 13, no. 1, pp. 42–52, 2021.

[19] W. Bekri, R. Jmal, and L. Chaari Fourati, "Internet of things management based on software defined networking: a survey," *International Journal of Wireless Information Networks*, vol. 27, no. 3, pp. 385–410, 2020.

[20] D. G. S. Pivoto, L. F. F. de Almeida, R. da Rosa Righi, J. J. P. C. Rodrigues, A. B. Lugli, and A. M. Alberti, "Cyber-physical systems architectures for industrial internet of things applications in industry 4.0: a literature review," *Journal of Manufacturing Systems*, vol. 58, pp. 176–192, 2021.

[21] G. Secundo, S. M. Riad Shams, and F. Nucci, "Digital technologies and collective intelligence for healthcare ecosystem: optimizing internet of things adoption for pandemic management," *Journal of Business Research*, vol. 131, pp. 563–572, 2021.

[22] T. Li, J. Yuan, and M. Torlak, "Network throughput optimization for random access narrowband cognitive radio internet of things (NB-CR-IoT)," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1436–1448, 2018.

[23] D. Kiel, C. Arnold, and K.-I. Voigt, "The influence of the Industrial internet of things on business models of established manufacturing companies-a business level perspective," *Technovation*, vol. 68, pp. 4–19, 2017.

[24] S. Wijethilaka and M. Liyanage, "Survey on network slicing for internet of things realization in 5G networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 957–994, 2021.

[25] T. Xia, M. M. Wang, J. Zhang, and L. Wang, "Maritime internet of things: challenges and solutions," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 188–196, 2020.

Research Article

A Study of Two-Way Short- and Long-Term Memory Network Intelligent Computing IoT Model-Assisted Home Education Attention Mechanism

Suling Ma 🆔

Lishui University, Lishui, Zhejiang 323000, China

Correspondence should be addressed to Suling Ma; masuling@sqgxy.edu.cn

This paper analyzes and collates the research on traditional homeschooling attention mechanism and homeschooling attention mechanism based on two-way short- and long-term memory network intelligent computing IoT model and finds the superiority of two-way short- and long-term memory network intelligent computing IoT model. The two-way short- and long-term memory network intelligent computing IoT model is improved and an improved deep neural network intelligent computing IoT is proposed, and the improved method is verified based on discrete signal homeschooling classification experiments, followed by focusing on the application research of the two-way short- and long-term memory network intelligent computing IoT model-assisted homeschooling attention mechanism. Learning based on neural network, human behavior recognition method combining spatiotemporal networks, a homeschooling method integrating bidirectional short- and long-term memory networks and attention mechanisms is designed. The visual attention mechanism is used to add weight information to the deep visual features extracted by the convolutional neural network, and a new feature sequence incorporating salient attention weights is output. This feature sequence is then decoded using an IndRNN independent recurrent neural network to finally classify and decide on the homeschooling category. Experiments on the UCF101 dataset demonstrate that the incorporation of the attention mechanism can improve the ability of the network to classify. The attention mechanism can help the intelligent computing IoT model discover key features, and the self-attention mechanism can effectively capture the internal features of homeschooling and optimize the feature vector. We propose the strategy of combining the self-attention mechanism with a bidirectional short- and long-term memory network to solve the family education classification problem and experimentally verify that the intelligent computing IoT model combined with the self-attention mechanism can more easily capture the interdependent features in family education, which can effectively solve the family education problem and further improve the family education classification accuracy.

1. Introduction

With the continuous improvement of the material level, the state's investment in family education has gradually increased, and people's demand for scientific family education concepts has become more and more urgent. The establishment of the National Family Education Association has provided favorable theoretical support for the development of family business in our country. Research articles and monographs on family education have emerged as the times require. Its research involves multiple perspectives and can be divided into three categories: one is popular science: reading materials; the second category is popular reading materials with a story complex; the third category is theoretical research works [1]. Family education is emotional, and an important feature of family education is that it is contagious. Family education is the communication between parents and children and the positive or negative influence of parents' words, actions, and behavior on girls [2]. Parents are their children's role models and life mentors in the process of growth. In the current situation, deep learning has achieved an overwhelming victory in the fields of speech recognition, natural language processing, and homeschooling recognition, which has greatly contributed to the

changes in other fields [3]. Applying deep learning to home education data processing can replace the traditional method of manually labeling data to achieve real-time processing of each piece of data received, which greatly saves the time of manual labeling and can effectively improve the efficiency of home education analysis [4].

Internet of Things technology, as a new generation of information technology, is an important third type of information technology after computer technology and Internet technology, and it is also a new wave ushered in by the information industry. The emergence of the Internet of Things regulates the communication and interactive control of information between people and things and things and things, which are integrated with the Internet through various wireless or wired networks. The attention model is applied to the analysis of the attention mechanism of family education, and the attention mechanism is added to the family education framework to weigh the input features to measure the importance of each feature to the current recognition, focusing on the important features and ignoring the unimportant features. The attention model is applied to family education attention mechanism analysis to extract feature information [5]. The self-attention mechanism, as a special kind of attention mechanism, is introduced into the homeschooling classification task to be able to learn the word dependencies within sentences and capture the internal structure of homeschooling. Therefore, this paper explores the current technical points in the field of family education analysis from the practical needs of family education analysis and seeks to find an efficient and accurate family education analysis method based on deep learning technology and self-attention mechanism. With the rapid development of deep learning technology, neural network models represented by convolutional neural networks and recurrent neural networks have obtained remarkable research results in the field of natural language processing, and people use these models for family education data analysis to obtain certain practical value [6].

While neural networks have so many advantages, they also inevitably have some drawbacks. The structure of neural networks and the corresponding training methods need to be specified in advance, which requires extensive design experience. The reasonable design of the network structure is very important to the accuracy of the modeling, the network structure is not reasonable to a large extent affecting the subsequent application of the network, and only the model structure is selected properly to make it play the proper value [7]. The selection of the initial value of the network also has an impact on the modeling, which is directly related to the convergence process of the weight training. If the initial value is not selected appropriately, it will prolong the training time, reduce the network accuracy, and weaken the generalization ability. Section 1 is the introduction. This section introduces the research background and significance of the paper and also outlines the organization of the paper. Section 2 is the related work. The direction and value of this paper's research are analyzed in light of the current research status at home and abroad. Section 3 is a study on the mechanism of home education attention based on two-way

short- and long-term memory network intelligent computing IoT model assistance. This section uses a combination of a two-way long- and short-term memory network and intelligent computing IoT model to classify the supervised education dataset, and the algorithm can greatly improve the classification accuracy of home education compared with the traditional method. Section 4 is the result analysis. This section analyzes the data for the classified family education dataset using a bidirectional short- and long-term memory network intelligent computing IoT model and attention mechanism techniques to achieve good family education classification results. Section 5 is the conclusion. This section systematically summarizes the research content of the thesis, analyzes the shortcomings, and provides an outlook on the next improvement direction.

2. Current Status of Research

The basis of computational intelligence can also be understood as the continuous evolution of structures. Since the birth of intelligent computing, its powerful intelligence, robustness, and self-adaptability have attracted the attention of many researchers, and great breakthroughs have been made in both theoretical research and practical applications of algorithms [8]. Kolomvatsos and Anagnostopoulos proposed a simple two-way short- and long-term memory network intelligent computing IoT algorithm for recommendation classification or nonrecommendation, which predicts the outcome by the average semantic of specific phrases in home education, with an average accuracy of 74% in the final experiment [9]. George and Santra proposed a combination of natural language processing and fuzzy logic techniques for analyzing homeschool content in homeschooling, and the experiments showed a very good correspondence between sentiment sets and human judgments of homeschool content [10]. Mollah et al. obtained rich emotional features by fusing machine learning methods and multiple rules and integrated them into the basic feature model, which improved the classification effect in the microblog sentiment classification experiment. Semantic rules are integrated into the support vector machine model to complete sentiment analysis tasks [11]. Experiments have verified that the support vector machine model combined with semantic rules is more effective in sentiment classification tasks [12]. The neural network can process the sequence information of text data and calculate and sort the influence of the input object on the final classification result. According to the sorting result, the words with strong emotional tendencies are assigned higher weights to reduce the loss of emotional information.

While neural networks have so many advantages, they inevitably have some shortcomings. The structure of the neural network and the corresponding training method need to be specified in advance, which requires us to have rich design experience [13]. Zhang et al. obtained C&W models by deep neural network training and applied them to processing tasks such as home education labeling and home education classification and achieved good results [14]. Ge et al. used convolutional neural networks for homeschooling

analysis tasks and conducted comparative experiments on seven datasets by combining convolutional kernels of different sizes with maximum pooling to demonstrate the effectiveness of single-layer CNNs in homeschooling classification tasks [15]. Wang and Ashwini proposed several recurrent neural networks such as RNN, MRNN, and RNTN, among which the RNTN model used syntactic analysis trees to obtain [16]. To imitate the sentence structure, McCormack et al. used a tree LSTM model to model family education and obtained excellent results in the family education classification task, which was input to the LSTM according to the temporal relationship to construct a CNN-LSTM model and apply it to the family education analysis task [17]. The content of family education should be rich and diversified, not limited to the traditional family education content, but diversified mainly in that the content of family education should involve all aspects of moral, intellectual, physical, aesthetic, and social development [18].

At present, how to apply the two-way short- and long-term memory network intelligent computing IoT model and attention mechanism to home education polarity analysis is still in the research stage, and the research goal of this thesis is to construct an efficient and accurate home education classification method. Based on this, the main research contents and innovations of this research work include the following three aspects. The loss function in deep neural networks has a significant influence on model training overfitting. To make the home education binary classification model more effective in overfitting the samples with prediction errors [19], local linearization in time limits the applicability of the proposed method in the discrete-time domain and impairs its prediction accuracy under strong nonlinear conditions. In the testing process, the idea of online learning was used for reference, and artificial judgment and relearning mechanisms were added to further improve the accuracy of network labeling. Experiments show that the automatic labeling algorithm can effectively improve the efficiency and accuracy of thermal measurement data analysis. According to the ranking results, higher weights are assigned to the categories with stronger family education tendencies to reduce the loss of family education information. Accordingly, this paper designs a bidirectional short- and long-term memory network intelligent computing IoT model and verifies the effectiveness of the model through quantitative and qualitative experiments on different homeschooling datasets [20].

3. Research on the Mechanism of Home Education Attention Based on Two-Way Short- and Long-Term Memory Network Intelligent Computing IoT Model Assistance

3.1. Feature Extraction of Family Education Attention Mechanism Based on Two-Way Long- and Short-Term Memory Network. In this paper, we combine BN-Inception with IndRNN to propose a deep learning human behavior recognition model with fused spatiotemporal networks. The spatial flow network deals with homeschooling attention

mechanism features, which are mainly used to capture features with strong differentiation for action understanding; the temporal flow network deals with continuous optical flow features, which are used to learn effective action features. This makes full use of the temporal and spatial dimensional information of the data, increases the depth of the network while reducing the complexity of the network, and finally achieves the purpose of improving the accuracy of behavior recognition. The feature model of the homeschooling attention mechanism based on the bidirectional short- and long-term memory network proposed in this paper is shown in Figure 1.

Since the obtained family education data also records the starting position of the effective value in the whole curve, firstly, the whole heat flow curve is divided into three sections according to the starting position, including the invalid data section before the effective data section, the valid heat flow segment, and the invalid data segment after the valid data segment. After that, divide the two large invalid data segments into several segments. The specific method is to use the length of the valid segment as the length of the data segment, and then from the beginning of the invalid data segment, follow the sliding window in steps of 0.9 times the length. The method moves backward in turn, thus dividing the invalid data segment into multiple segments. This article adopts the variation equation (1) of the family education problem. The first term is the regularization term to obtain a smooth displacement field.

$$
\text{variation} = \min\Bigl(\int \Bigl| \forall (x_1)^2 + \forall (x_2)^2 \Bigr| \mathrm{d}\varpi \\
\cdot \int \Bigl| H_1 (x + u(x)) - H_0 (x)^2 \Bigr| \varpi \Bigr).
\tag{1}
$$

The mean intensity difference between pixels is therefore used as the homeschooling similarity score. Thus, the target parallax map is the minimum of equation (2). The optical flow algorithm uses the H_1 parametrization to solve for minimization of the error function and minimization of the regular term of the target homeschool. Given two homeschooling frames H_0 and H_1, the objective is to find the parallax map such that the error function and the canonical term of homeschooling are minimized simultaneously, where the first term is the fidelity of the homeschooling data and the second term is the a priori canonical term. β denotes the correlation coefficient between the fidelity and the canonical term.

$$
G(x) = \int \Bigl[\alpha \beta (H_0 (x) - H_1 (x + u(x))) + \varpi (x_1, x_2) \Bigr] \mathrm{d}x.
\tag{2}
$$

The effect of education depends on the consistency of the influence of the school and the family. If there is no such consistency, then the school's teaching and education process collapse like a paper house. School teachers communicate with parents in time through home visits and parent meetings. The weight coefficient of each subsection is learned through the visual attention mechanism, and the probability distribution of each subsection is finally fused,
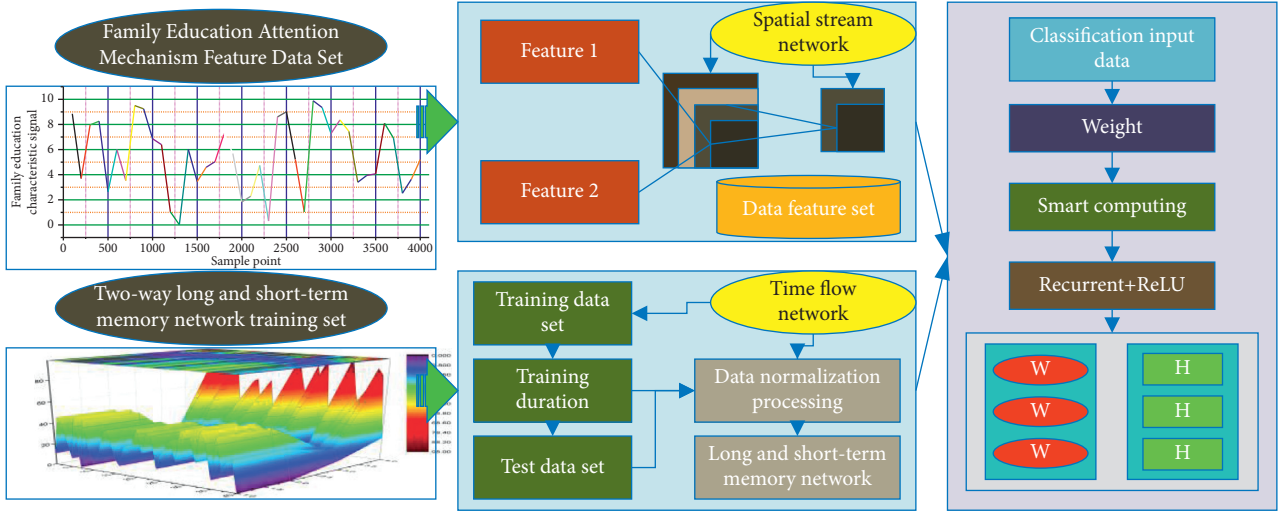
Figure 1: The characteristic model of family education attention mechanism.

and the IndRNN cyclic neural network is input for sequence learning. Finally, the results of IndRNN are used in the behavior prediction of the family education level.

The bidirectional short- and long-term memory network considers the temporal dimensional association of the convolutional neural network feature sequences, and the pooling operation only produces feature vectors that keep the temporal order constant. We assume that the feature sequences extracted by the spatial and temporal flow networks are $X = (x_1, x_2, \ldots, x_T)$, so the independent recurrent neural network (IndRNN) computes a sequence of hidden vectors with N neurons at each moment t $H = (h_1, h_2, \ldots, h_T)$ and the output vector sequence $f(t) = (y_1, y_2, \ldots, y_T)$, satisfying equation (3), where A denotes the weight coefficient from the previously hidden layer to the current hidden layer, and $A(h_0)$ denotes the weight coefficient from the hidden layer to the output layer, β denotes the deviation coefficient, and $h(t)$ denotes the hidden vector of the nth neural unit at moment t.

$$\begin{cases} n = [1, 2, 3, \ldots, N], \\ f(t) = A(h_0) * h_1 + \beta(t), \\ h(t) = H\left(A(h_i)\infty * y(t) + A(h_h) * y(t-1) + \beta(t)\right). \end{cases} \tag{3}$$

The computation of the bidirectional LSTM model in the hidden layer is the same as that of the hidden layer of the LSTM and is passed between the same layers and the parameters are shared. Therefore, let the parameters of forwarding propagation be V and the parameters of backward propagation be U. Then the forward propagation formula of the bidirectional LSTM is as follows:

$$\begin{cases} \text{Lstm}(h_t) = B(t) * \sin(h(\zeta(t))), \\ \text{Lstm}(h_t^*) = B^*(t) * \sin(h(\zeta(t)^*)). \end{cases} \tag{4}$$

After the forward propagation of the network is finished, the output is done as a Softmax operation to obtain the predicted value $\hat{y}$ with the true label value y to calculate the

loss function. The experiment uses cross-entropy as the loss function, so the derivation of the backpropagation process of the bidirectional LSTM is the same as the derivation of the backpropagation process of the LSTM, as in the following equation:

$$\frac{\mathrm{d}(f(t))}{\mathrm{d}W} = \beta * (1 - \beta) * (\beta(t) + +\zeta(t) \\ + B^* * f(t) + W * M(t-1)^*). \tag{5}$$

3.2. Intelligent Computing IoT Model Construction. The practical activities in the virtual-real integration environment are different from the traditional practical activities, which put forward higher requirements on teaching, environment, model, and process. Based on the characteristics of the comprehensive practical activities course and the current situation research, this study, combined with the Internet, the characteristics of the IoT, and other information technology, made adaptable modifications based on Cooper's experiential cycle learning circle and summarized the comprehensive practical activities teaching based on experiential learning theory. The five elements of the model are precursor learning, practical inquiry, extension and expansion, observation and reflection, and communication and evaluation. The calculation process of the self-attentive mechanism is divided into three steps. Step one is to calculate the similarity between an element and a keyword, the second step is to normalize the original score, and finally, the weighted sum of the weighted values according to the obtained weighting coefficients is used to find the value of self-attentive for the element. The keyword is denoted by K, the function is denoted by F, the query is denoted by Q, the data source is denoted by S, the similarity is denoted by Sam, the attention value is denoted by A, and the value of the weight coefficient is denoted by V. Here $Q = K = V$; when home education is entered, self-attention is calculated once.

Different functions and computer mechanisms are introduced according to the query and keywords, and then the similarity between the two is calculated, and the intelligent calculation of the IoT calculation formula is shown in the following equation:

$$\text{Sam}(Q, K) = \sum_{i=1}^{N} MLP(Q, K_i). \tag{6}$$

A calculation method similar to Softmax is introduced to convert the score obtained in the previous step to obtain the weight coefficient ai. The calculation formula is formula (7). Determine a confidence threshold above the threshold is a valid data segment, and below the threshold is an invalid data segment. Input the multiple pictures finally obtained in the third step into the trained convolutional neural network again to obtain the confidence level of each picture. These confidences are compared with the previous confidence thresholds, and two demarcation segments are obtained, which are lower than the threshold and higher than the threshold.

$$f(x)\mathbb{F}ax(\text{Sam}(x)) = \frac{x^{\text{Sam}(i)}}{\sum_{i=1}^{N} x^{\text{Sam}(i)}}. \tag{7}$$

When the input sequence n is smaller than the representation dimension d, the time complexity self-attention mechanism of each layer is the smallest. When n is relatively large, one solution is a restricted self-attention mechanism; that is, each word is not calculated with all words, but only with restricted r words. The force mechanism and the convolutional neural network can be well paralleled and do not rely on the previous calculation, which is better than the cyclic neural network; in terms of long-distance dependence, the self-attention mechanism can combine any two sentences in a sentence through a calculation result. Words are connected, so the long-distance dependence feature is greatly shortened, and its maximum path length is 1, which can capture long-distance dependence. The text data within a certain sliding window is converted into a fixed-length vector B'' by the Softmax function, and then the feature data is fed into a linear SVM classifier for training and outputting the sentiment classification results, which is calculated by equation (8), where C denotes the training error of the SVM classifier, W and H denote the final obtained weight matrix and bias value, respectively, and L denotes the loss function.

$$\begin{cases} B'' = \text{Soft max}(w * Bi + \beta), \\ \\ C(W, M) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} H_{ij}(B'')}{M * N} + \alpha * Y(W). \end{cases} \tag{8}$$

When the neural network model is trained, the selection and setting of parameters is a very important aspect, and the training results will be very different when different parameters are selected, and the parameter settings of the model in this section are shown in Table 1.

Table 1: Model parameter configuration.

Serial number	Name	Meaning
1	Batch size	Number of batch training samples
2	Learning rate	E-learning efficiency
3	VALIDATION_SPLIT	The proportion of test set in training data set
4	Return sequences	Network return type
5	Early stopping	Model stop training parameters
6	NB epoch	Model iteration times

The cross-entropy loss function, also called the logarithmic loss function, is designed for the performance comparison of classification models and can be divided into dichotomous cross-entropy and multiclassification cross-entropy according to the difference of whether the classification model is dichotomous or multiclassification. It has the mathematical expression of the following equation:

$$G(x) = \sum_{i=1}^{N} \sum_{j=1}^{N} g(x_i) + \log Y(x_i), \quad x_i = y(i) + j. \tag{9}$$

Traversing the entire dataset to compute the loss function once and update the gradient of each parameter, this method is called batch gradient descent, and the parameters are updated using equation (10). $G(x)$ is the value of the kth parameter at x iterations, $H(x)$ is the first-order partial differential of the error function corresponding to that parameter, and ρ is the step size (learning rate), which indicates the size of the step, which determines the length of each step along the negative direction of the gradient, and a current optimal value is generally obtained by the linear search, which increases the computational effort and affects the performance of the model and does not support online learning.

$$G(x + 1) = G(x) - \rho * \forall G(x) * H(x). \tag{10}$$

3.3. Classification Evaluation of Family Education Attention Mechanism.

When a predictive classifier is used to process binary sample data, we usually classify the sample data into two categories: Positive (Positive) and Negative (Negative). When the family education sample is classified as positive, it means the classification is correct, so we call it True Positive (PT) or true positive; when the positive family education sample is classified into the negative family education sample, it means the classification is wrong. Call it False Negative (NF) or false negative. Similarly, when dealing with negative sample data, if the classifier classifies the negative sample data into a negative class, the negative sample data is correctly classified, and we call it True Negative (NT) or true negative; if the negative sample data is classified into positive class, it means the classification is wrong, and we call it False Positive (FP) or false positive. The prediction effect is evaluated according to the value of the metric in the following equation:

$$
\begin{cases}
\text{NS} = \dfrac{\text{PT}}{\text{NF} + \text{PT}} * 100\%, \\[2ex]
\text{PS} = \dfrac{\text{NT}}{\text{NT} + \text{PF}} * 100\%, \\[2ex]
\text{ACC} = \dfrac{\text{NT} + \text{PT}}{\text{TOTAL}} * 100\%, \\[2ex]
\text{TOTAL} = \text{PF} + \text{PT} + \text{NF} + \text{NT}, \\[2ex]
\text{MCC} = \dfrac{(\text{PT} * \text{NT}) - (\text{PF} * \text{NF})}{\sum \sqrt{(\text{PT} + \text{PF}) * (\text{PT} + \text{NF}) * (\text{NT} + \text{PF}) * (\text{NT} + \text{FNF})}} * 100\%,
\end{cases}
\tag{11}
$$

where NS is the sensitivity metric for classification, and the magnitude of the NS value can be observed to understand how well positive samples are classified as positive samples. PS is the classifier specificity metric, which calculates the percentage of negative sample data that are classified as negative classes. Acc is the accuracy rate, and the metric reflects the percentage of positive and negative samples that are both correctly classified at the time of classification. MCC value, also known as the MCC correlation coefficient metric, reflects the Sn and PS metrics between the equilibrium degree of NS and PS is shown by the above equation; when the difference between these two indicators is smaller, MCC is larger.

In the evaluation process of the attention mechanism of family education, there are soft evaluation aspects such as the humanization of model construction and design and the innovation of model construction. Therefore, the evaluation method of this article adopts very good, good, average, poor, and very poor. Describe the related problems, realize the primary level evaluation, and then use fuzzy mathematics to make the fuzzy information numerical, to carry out quantitative analysis and evaluation.

When multiple features are included in the training model, the dimensionality of the mixed feature vector may become very large, and due to the presence of more random information and noise in the uneven initial features, it will hurt the training model, and usually to solve such problems a feature selection tool such as F-score [25] can be used to select the most effective features. Therefore, we also call F-score a metric to identify features between different categories based on the F-score value, and the larger the F-score value, the stronger the differentiation of such features between different categories. The expressions for calculating F-score are shown in the following equation:

$$
F\,(\text{score}) = \frac{\sum_{i=1}^{M}\left(x_i^+ - x_i^-\right)}{m^+ - 1} + \frac{\sum_{i=1}^{M}\left(x_i^+ - x_i^-\right)}{m^- - 1}.
\tag{12}
$$

In some of the discrete classifiers or two-classified classifiers we use, their classification decisions (Y or N) at each output can be represented by a single point in the ROC space, and all the points in the threshold interval converge into a curve, the ROC curve, so we can use this ROC curve to represent this threshold interval. Further, we express the performance depicted by the ROC curve, the AUC value, with the formula as in equation (13), while the specific diagnostic descriptions are shown in Table 2.

$$
\text{Value}\,(\text{AUC}) = \int_{-\infty}^{+\infty} (NS(x) * \text{FPR}(x))\mathrm{d}x.
\tag{13}
$$

In equation (13), x denotes the cutoff value of classification prediction probability and FPR denotes the percentage of negative samples being predicted as positive samples for the following equation:

$$
\text{FPR} = \frac{\text{PF}}{\text{NT} + \text{PF}}.
\tag{14}
$$

4. Analysis of Results

4.1. Analysis of Bidirectional Short- and Long-Term Memory Networks. The introduction of an attention mechanism can control the influence of elements in the input sequence on the elements of the output sequence and retain more important information; based on the CNN-BO and LSTM-BO models, the M value is selected from 0.1 to 1.0, increasing by 0.1, respectively. The experimental results are shown in Figure 2. The accuracy of the two models increases and then decreases with the value of N. The accuracy of the LSTM-BO model is 82.25% and that of the CNN-BO model is 88.74% at $M = 0.5$, which are both the maximum values. The loss rate of the two models shows a general trend of decreasing and then increasing with the increase of M value, and at $M = 0.5$, the loss rate of the LSTM-BO model is 0.237 and the loss rate of the CNN-BO model is 0.167, which are both the minimum values. The above experimental results are analyzed, and the threshold M of the two models in this paper is taken as 0.6.

Dropout is an optimization method proposed to solve the overfitting and gradient disappearance problems of deep neural networks. To investigate the effect of dropout on the training results, this experiment takes the value of dropout from 0.1 to 1.0 and increases it by 0.1 in order, and the experimental results are shown in Figure 3. The highest accuracy of the LSTM-BO model is 79.86%, the lowest loss

TABLE 2: Algorithm performance diagnosis description.

Serial number	AUC value range	Forecast description
1	AUC < 0.5	Not in line with the real situation
2	0.5 ≤ AUC < 0.8	Low accuracy
3	0.8 ≤ AUC < 0.8	Precise
4	AUC ≥ 0.8	With high accuracy

(a)

(b)
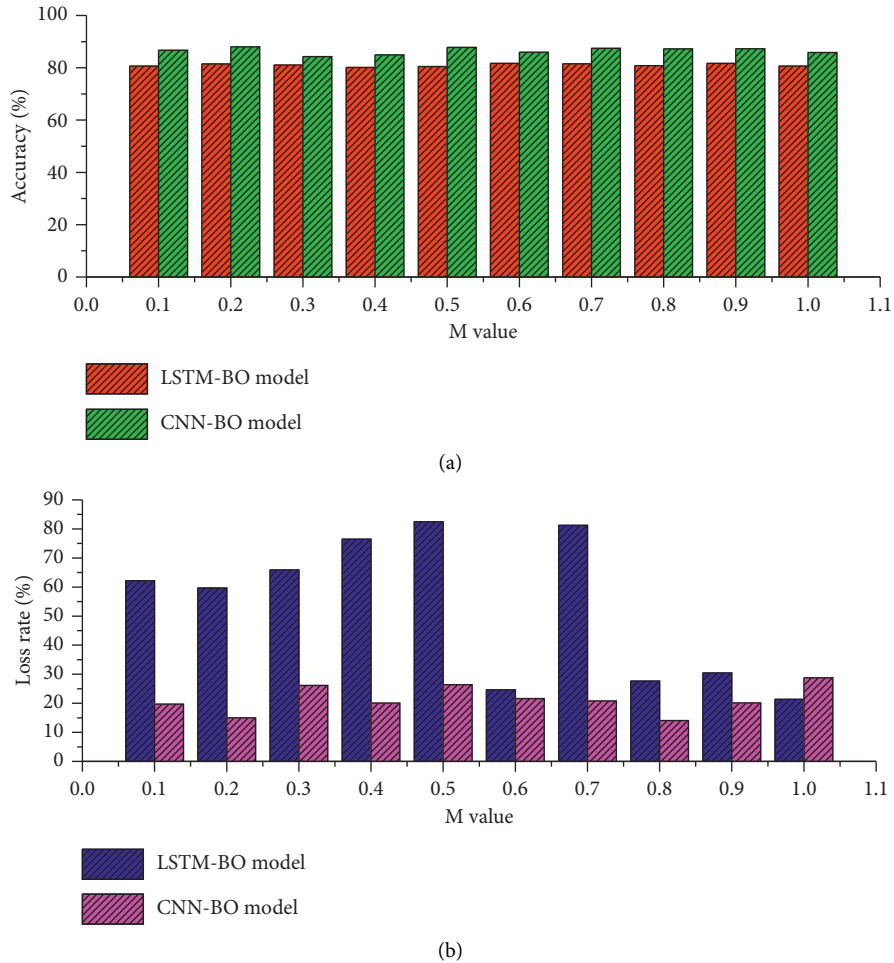
FIGURE 2: Accuracy and loss rate variation.

rate is 0.237, and the lowest time consumption is when the dropout is set to 0.2. The highest accuracy of the CNN-BO model is 87.75%.

4.2. Intelligent Computing IoT Model Analysis. The accuracy and F-value of the intelligent computing IoT model in the training process are shown in Figure 4, where Acc, measure, and loss in Figure 4 denote the accuracy and F-value of the training set in the training process, and val_acc and val_fmeasure denote the accuracy and F-value of the validation set in the training process, respectively. From Figure 4, it can be seen that, after 50 layers of iterations, the change curves of the four indicators have leveled off, indicating that the model in this chapter has reached the best training effect; from Figure 4, it can be seen that, after 50 layers of iterations, the loss in training has been reduced to

the lowest level, indicating that the training goal of the intelligent computing IoT model has been achieved.

The final experimental results of the four groups of models are shown in Figure 5. From the data in Figure 4, it can be seen that the algorithm model proposed in this chapter works best, and compared with the traditional CNN model, the accuracy of this algorithm increases from 87.63% to 92.51%, and the F-value increases from 87.92% to 93.22%, while its algorithmic performance far exceeds that of the SVM algorithm, which performs relatively well among the traditional machine learning algorithms.

4.3. Evaluation Analysis of Family Education Attention Mechanism. We converted the performance depicted by the ROC curve into a single scalar value by calculating the area under the ROC curve. Of course, the AUC serves as a more

(a)

(b)

(c)
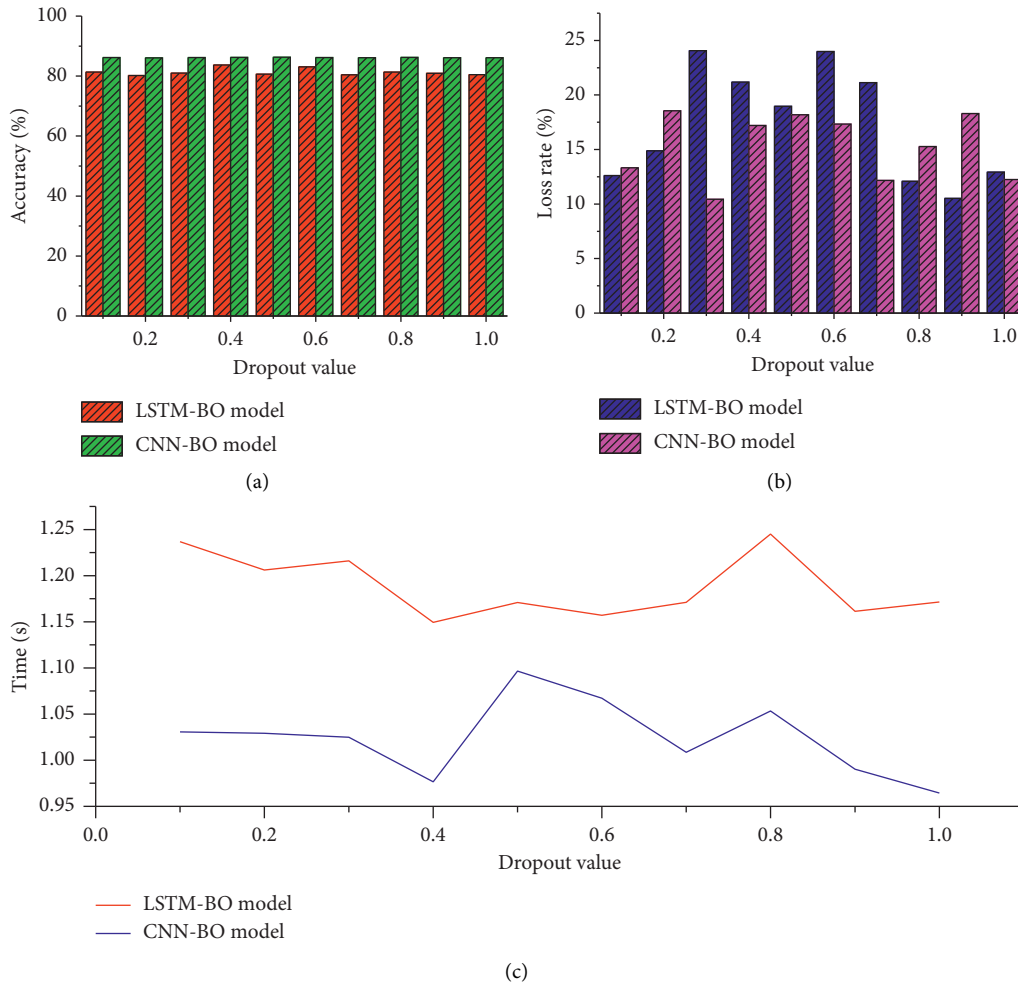
Figure 3: Dropout selection experimental results.
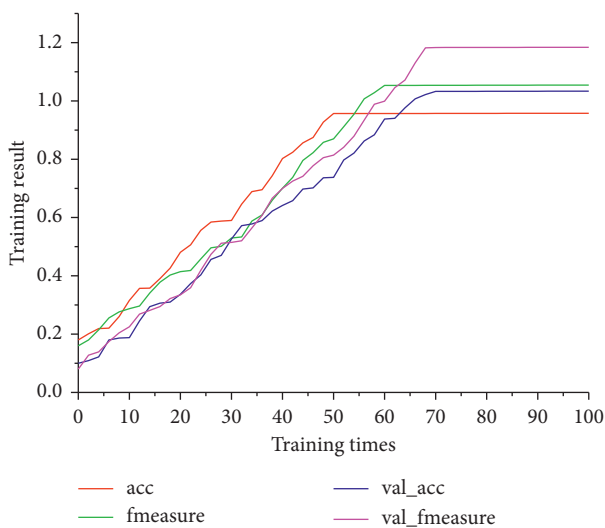
Figure 4: Accuracy and F-value in training.

popular binary classifier evaluation metric, where a larger AUC value indicates better performance in the prediction of family education attention mechanisms. In Figure 6, the blue

curve is generated from the first layer of the family education attention mechanism prediction, while the red curve is generated from the second layer of the family education attention mechanism prediction. The corresponding AUC values can also be seen in Figure 6. The AUC values for the predicted enhancer and enhancer strength categories are 0.934 and 0.842, respectively, and it is clear that, for identifying enhancers and nonenhancers, this AUC value is higher than the AUC value for the predicted enhancer strength category. However, both values reached a high level, which shows that the prediction of the family education attention mechanism we constructed has some accuracy and stability. When developing a classifier for predicting the attention mechanism of family education, the most important thing is to achieve the highest possible sensitivity while maintaining a low false alarm rate. The advantage of the ROC curve is that you can visualize all these performance indicators from a single image. Comparing the curve with the competitive model is a quick and easy way to choose the appropriate classification or diagnostic tool.

The dataset we used was of the same standard as the datasets of several other studies, and after we completed our prediction study to some extent, we compared the

(a)

(b)

(c)

(d)
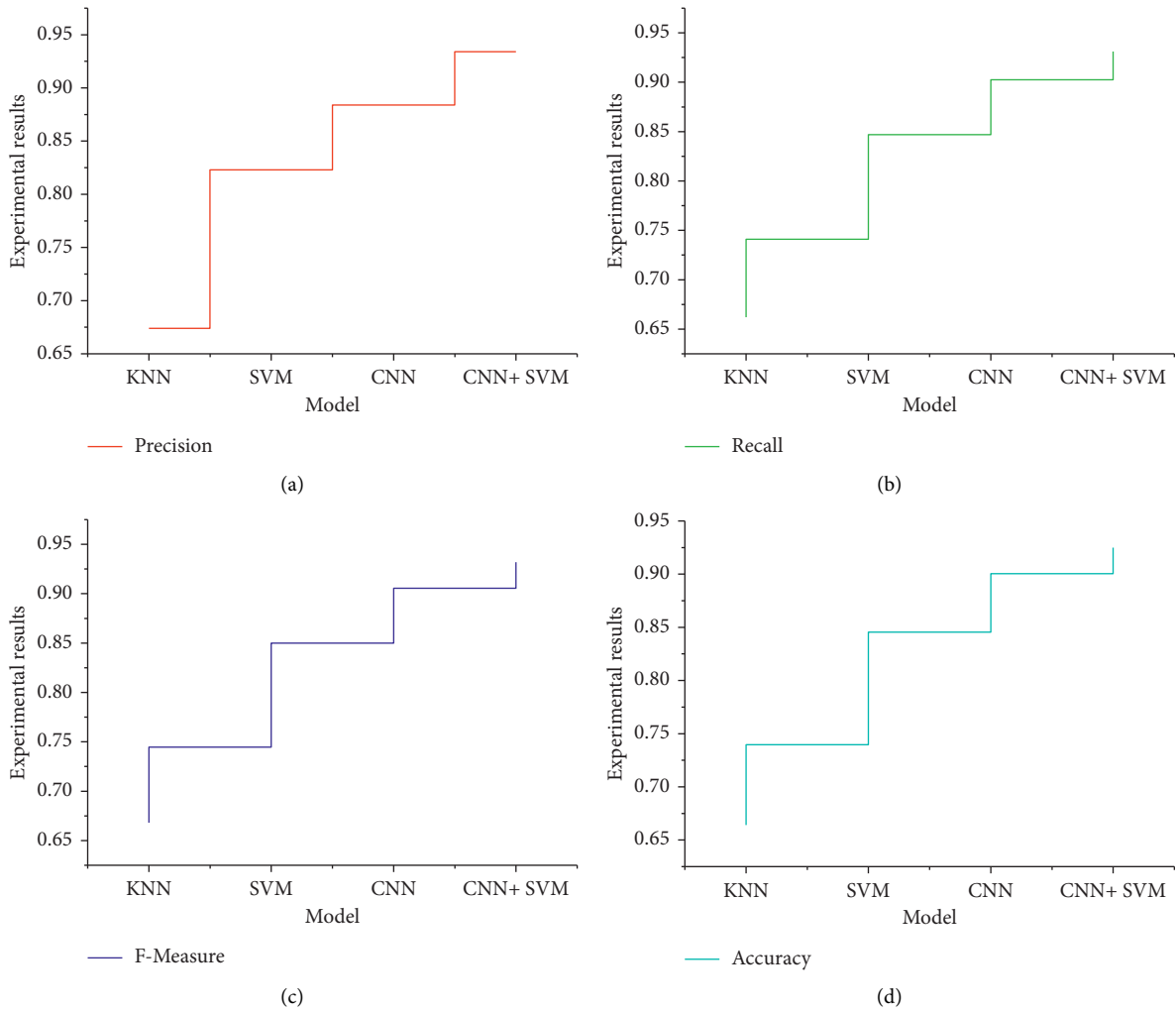
Figure 5: Experimental results.
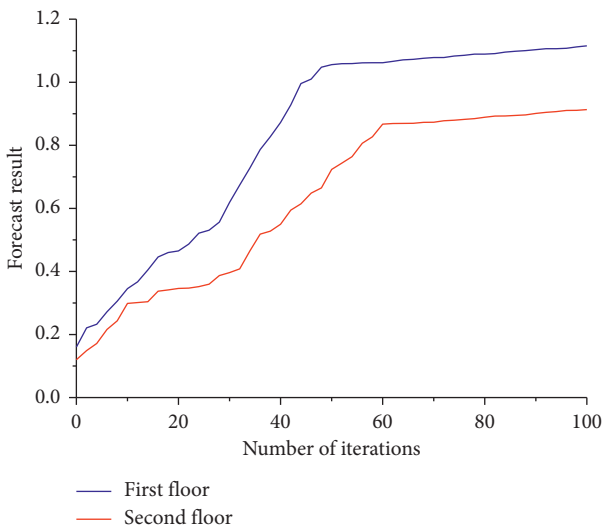
Figure 6: The ROC plot of the predicted attention mechanism of family education.

classification prediction results of various research methods to be able to analyze the effects of the study more objectively, and the results are shown in Figure 7. Comparing the Acc values of various family education attention mechanism predictions in the table, we can see that our constructed family education attention mechanism prediction iESW-2L2 is better than several existing family education attention mechanism predictions in identifying enhancers and non-enhancers, and the MCC value of the first layer of this family education attention mechanism prediction is the highest, indicating that this family education attention mechanism prediction has relatively good stability, and this paper's homeschooling attention mechanism prediction has generally improved in the first tier structure of the study. In addition, we can see from Figure 7 that the Acc value in the second layer of the family education attention mechanism prediction reaches 66.22%, which is second only to iEnhancerPred's 68.21% but has the best stability, indicating that the study has some effectiveness.
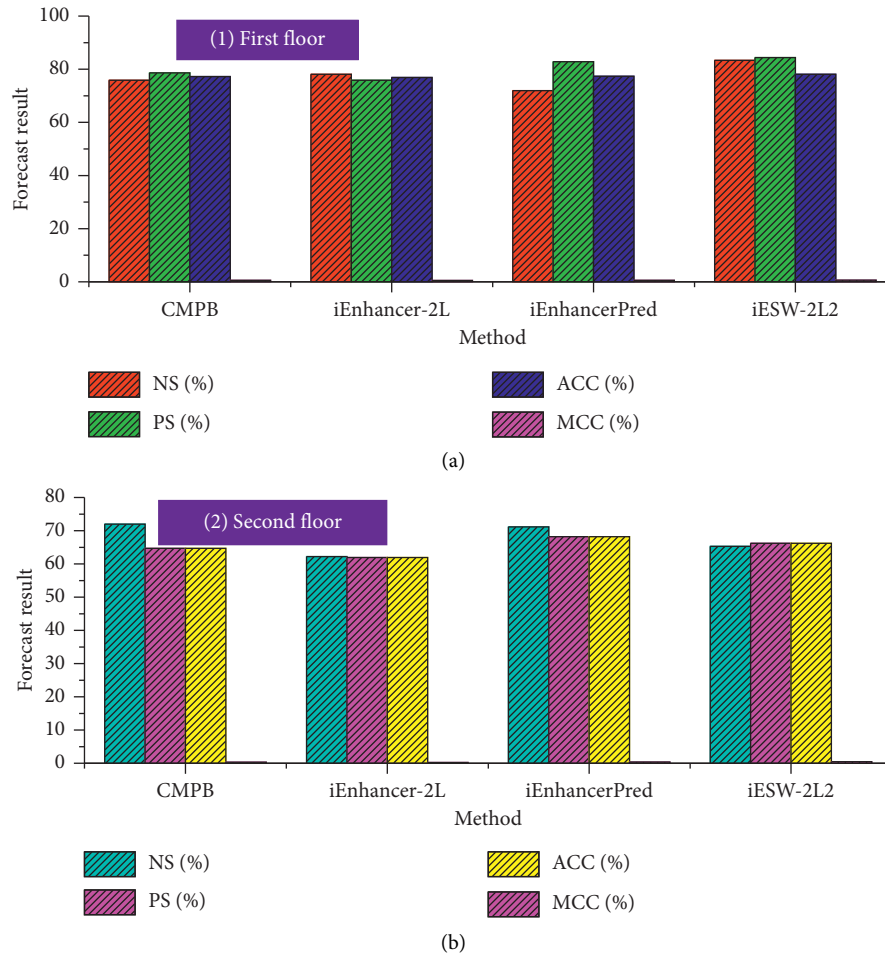
FIGURE 7: Comparison of prediction results on the benchmark dataset.

5. Conclusion

Firstly, the current mainstream model deep neural network methods inside the field of artificial intelligence and big data research, forward propagation, and backward feedback methods in deep neural networks, classification of neural networks, and the structure and application of short- and long-term memory neural networks in recurrent neural networks are studied and analyzed. On this basis, the hidden layer activation function of the two-way recurrent neural network for short- and long-term memory is improved, and the improved neural network model is proposed in combination with the Morlet wavelet function. Based on simple discrete single-symbol text classification experiments, the effectiveness and performance of the model proposed in this paper are verified, by briefly elucidating the principles of the original recurrent neural network and the short- and long-term memory neural network and proposing the combination of the traditional wavelet function with the short- and long-term memory neural network, as well as changing the model levels and activation functions to improve the experimental results. The core goal of the attention mechanism is to quickly filter out high-value information from a large amount of data.

The introduction of the attention mechanism in the homeschooling classification task can learn the dependencies within the court education and capture the internal structure of homeschooling. A two-way long- and short-term memory network intelligent computing IoT model to assist the attention mechanism of homeschooling is studied, a two-way long- and short-term memory network intelligent computing IoT model based on the two-way long- and short-term memory network is constructed, and its parameter modification rules of BP algorithm based on gradient descent are derived, respectively. To improve the recognition accuracy, a two-layer learning strategy is introduced, and the BP algorithm is used as the inner learning algorithm, while the outer learning adopts an improved particle swarm optimization algorithm, and the main innovation is to introduce the early convergence degree evaluation index to divide the population into three parts and adjust them adaptively to jump out of the local optimum, respectively. In this paper, we mainly study the sequence order information features by using various techniques to extract the sequence order information features. For feature extraction such as sequence structure, clustering algorithms, artificial intelligence analysis, and other methods can be used for in-depth research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

Acknowledgments

References

[1] Y. Song, S. Tian, and L. Yu, "A method for identifying local drug names in xinjiang based on BERT-BiLSTM-CRF," *Automatic Control and Computer Sciences*, vol. 54, no. 3, pp. 179–190, 2020.

[2] M. Yang, X. Li, and Y. Liu, "Sequence to point learning based on an attention neural network for nonintrusive load decomposition," *Electronics*, vol. 10, no. 14, pp. 1657–1721, 2021.

[3] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: a survey and new perspectives," *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, pp. 10–38, 2019.

[4] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2923–2960, 2018.

[5] J. Liu, H. Guo, J. Xiong, N. Kato, J. Zhang, and Y. Zhang, "Smart and resilient EV charging in SDN-enhanced vehicular edge computing networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 217–228, 2019.

[6] K. Hsiao, C. Noble, W. Pitman et al., "A thalamic orphan receptor drives variability in short-term memory," *Cell*, vol. 183, no. 2, pp. 522–536, 2020.

[7] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of things: a survey of emerging technologies," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1121–1167, 2020.

[8] C. M. Souza, M. R. G. Meireles, and P. E. M. Almeida, "A comparative study of abstractive and extractive summarization techniques to label subgroups on patent dataset," *Scientometrics*, vol. 126, no. 1, pp. 135–156, 2021.

[9] K. Kolomvatsos and C. Anagnostopoulos, "A probabilistic model for assigning queries at the edge," *Computing*, vol. 102, no. 4, pp. 865–892, 2020.

[10] S. George and A. K. Santra, "Traffic prediction using multifaceted techniques: a survey," *Wireless Personal Communications*, vol. 115, no. 2, pp. 1047–1106, 2020.

[11] M. B. Mollah, J. Zhao, D. Niyato et al., "Blockchain for future smart grid: a comprehensive survey," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 18–43, 2020.

[12] J. Joseph and A. J. Wilson, "The growth of the firm: an attention-based view," *Strategic Management Journal*, vol. 39, no. 6, pp. 1779–1800, 2018.

[13] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila, and T. Taleb, "Survey on multi-access edge computing for Internet of things realization," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2961–2991, 2018.

[14] R. Zhang, X. Jing, S. Wu, C. Jiang, J. Mu, and F. R. Yu, "Device-free wireless sensing for human detection: the deep learning perspective," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2517–2539, 2020.

[15] X. Ge, Q. L. Han, X. M. Zhang, L. Ding, and F. Yang, "Distributed event-triggered estimation over sensor networks: a survey," *IEEE Transactions on Cybernetics*, vol. 50, no. 3, pp. 1306–1320, 2019.

[16] S. Lalitha and M. V. Ashwini, "A study on artificial intelligence application in higher education institutions[J]," *Solid State Technology*, vol. 63, no. 6, pp. 21797–21808, 2020.

[17] J. McCormack, P. Hutchings, T. Gifford, M. Yee-King, M. T. Llano, and M. D'inverno, "Design considerations for real-time collaboration with creative artificial intelligence," *Organised Sound*, vol. 25, no. 1, pp. 41–52, 2020.

[18] A. E. Putra, K. Rukun, D. Irfan et al., "Designing and developing artificial intelligence applications troubleshooting computers as learning aids," *Asian Social Science and Humanities Research Journal (ASHREJ)*, vol. 2, no. 1, pp. 38–44, 2020.

[19] A. Alwarafy, K. A. Al-Thelaya, M. Abdallah, J. Schneider, and M. Hamdi, "A survey on security and privacy issues in edge-computing-assisted Internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4004–4022, 2020.

[20] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2671–2701, 2019.

Research Article

Model Construction and Research on Decision Support System for Education Management Based on Data Mining

Weidan Wang (ID)

School of Political and Public Management, Zhengzhou University, Zhengzhou, Henan 450000, China

Correspondence should be addressed to Weidan Wang; wendywang@gs.zzu.edu.cn

Based on data mining technology, this paper applies a combination of theoretical and practical approaches to systematically describe the background and basic concepts related to the generation of data mining-related technologies. The classical data mining process is analyzed in depth and in detail, and the method of building a decision support system for education management based on the B/S model is studied. Not only are the data mining techniques applied to this system, but also the decision tree model with the improved ID3 algorithm is implemented in this thesis, which is further applied to the educational management decision support system of this topic. The load of the client computer is reduced, and the client computer only needs to run a small part of the program. This paper focuses on the following aspects: the overall planning of the educational management decision support system based on data mining technology. From the actual educational management work, we analyze the requirements and design each functional module of this system in detail, applying the system functional structure diagram and functional use case diagram to represent the functional structure of the system and using flow charts to illustrate the workflow of the system as a whole and in parts. The logical structure design, entity-relationship design, and physical model design of the database have been carried out. To improve the efficiency of the system, the ID3 algorithm was improved on this basis to reduce the time complexity of its operation, improve the efficiency of the system operation, and achieve the goal of assessing and predicting the teaching quality of teachers. The development and design of this system provide an efficient, convenient, scientific, and reliable system tool to reduce the workload of education administrators and, more importantly, to make reasonable and effective use of the large amount of data generated in the management, and data mining techniques are used to extract valuable and potential information from these data, which can be more scientific and efficient for the teaching of teachers and students. It can provide reliable, referenceable, and valuable information for managers to make assessments and decisions.

1. Introduction

In the current context of rapid economic growth and rising urbanization, education resources, as one of the important social public resources, are lagging in terms of technical means of management, which poses a negative impact on the deepening of education reform and the optimal allocation of education resources, thus making the development of these tasks face serious challenges [1]. With the continuous progress of information technology, information technology has been widely used in the field of education management at home and abroad to improve their management level [2]. Although the educational information management systems currently in use are also based on database construction,

such systems can only achieve simple analysis and management of data, and the decision-making function is relatively weak. The system can collect education-related data and conduct simple analysis and statistics on the collected data, but its functions are relatively weak or even absent when it comes to the comprehensive analysis of education management. Many developed countries in the world have established a complete, high-tech, and scientific education management decision support system, which has a huge guiding role in the economic development of their national systems. Many colleges and universities have also established educational management decision support systems that are suitable for their own development in local area networks and wide area networks [3]. At the same time, the use of data

mining technology in education management decision-making has also received more and more attention from universities, and better application of data mining to provide services for education management and decision-making has become a new issue facing education [4].

For quite a long time in the past, educational decision-making was based on leadership intuition as well as relevant experience or even relied on social trends, which is obviously not in line with scientific decision theory, and therefore became an important criticism content for educational research staff. Scientific decision-making is inseparable from the necessary data support, but, for a long time in the field of education, a lot of data helpful for decision-making is often scattered on the desks and beds of faculty members [5]. Therefore, strengthening the construction of information technology systems, centralizing and unifying the management of large amounts of data, and then using the corresponding algorithms to analyze them will enable decision-making managers to quickly obtain a large amount of data support, helping them to more accurately understand the strengths and weaknesses of the current educational reality to obtain scientifically correct decisions. Data mining technology can analyze large and complex datasets efficiently and quickly and can find out the information hidden behind the data and the correlations, trends, and directions among the data [6]. If data mining technology is adopted in the education management system, it has important practical significance for the long-term development of education management. We are still continuing to study improved decision tree algorithms, and corresponding improvements have been made to the C4.5 algorithm from different perspectives. Among them are the time-consuming improvements for the C4.5 algorithm to process continuous attributes, using mathematical methods. The price is infinitesimal to improve the calculation efficiency of the information gain rate and so on [7].

The study conducted requirements analysis and detailed design of each functional module for this system and applied system functional structure diagram and functional use case diagram to describe the functional structure of the system and flow chart to illustrate the workflow of the system as a whole and in parts. The logical structure design, entity-relationship design, and physical model design were carried out for the database. Then the traditional ID3 algorithm was applied to the teaching quality assessment subsystem for analysis. To improve the efficiency of the system, the ID3 algorithm was optimized and improved on this basis to reduce the time complexity of its operation and improve the system operation efficiency. The goal of assessing and predicting teachers' teaching quality is achieved. This paper is divided into 5 sections, and the structure is arranged as follows: Section 2 discusses related work, the current state of research, and the application of classical algorithms, which lays the foundation for the research use in the subsequent chapters. Section 3 is devoted to research on data mining-based decision support system for educational management. Section 4 is devoted to results analysis. Section 5 gives the conclusion. Data mining techniques are applied to educational management decision support system, relevant

teacher information data, student information data, and social background information data are collected, and then the data are preprocessed; data mining algorithms are used to build models, test models, and apply models to help educators understand students' learning characteristics as much as possible, and the factors affecting teaching quality are analyzed. The value of the data related to students' learning information is maximized, and the role of educational data mining is truly brought into play.

2. Related Work

With the acceleration of campus digitalization in recent years, many universities have developed information management systems on their own or jointly with other universities to realize paperless offices and improve the scientific and high efficiency of daily work. The use of this system has accumulated a large amount of data, but colleges and universities have only made some inquiries, statistics, and reserves on these data and have not made full use of them. To make better use of these data and discover hidden valuable information, some universities have started to use data mining techniques to study, analyze, and solve problems. Therefore data mining technology will get more attention and application in the field of education [8]. Rodrigues et al. studied the application of data mining in the analysis of college entrance examination admission data through association rules and decision tree classification in the design of the college entrance examination information analysis system. Useful knowledge is mined for educational institutions, candidates, and colleges and universities [9]. Ghorbani and Ghousi conducted an in-depth study on educational management and data mining techniques in colleges and universities, using data mining related algorithms, and integrated a simple voting strategy through the study of simple improvement method of Apriori algorithm and ID3 algorithm, which has improved the educational management of students in colleges and universities and provided valuable guidance for students' training [10]. Farouk and Zhen used the educational technology of OLAM to analyze the education and admission of colleges and universities and identified the reasons and potentially useful information affecting students' enrolment through data mining techniques based on the educational data of previous years [11].

To pursue the rationality of educational decision-making and promote educational decision-making from experience to science, scholars at home and abroad have done a lot of research and discussion on educational decision support from different perspectives. Du et al. started the theoretical research in this area by proposing a data mining-based decision support system for educational management [12]. Shen et al. used the B/S model and analyzed the process dynamics of learners in the Moodle learning platform using network analysis techniques. They saved the learner dynamic record data in the management system to form real-time analytical models or statistical reports [13]. Operations such as classification patterns, clustering patterns, association rules, and other analytical methods as well as visual

representations were formerly used, and the latter were used to represent common model generation from behavioural dynamic data. Lee et al. used data mining techniques to estimate the potential of the subjects [14]. Based on the organic integration of information about learners' meta-cognition, motivational goals, knowledge acquisition, and learning attitudes, they established models with learners as the objects and predicted the future development tendencies of learners, analyzed and improved educational method models, explored the effectiveness of various applications to learning systems, and established data-based computing models to improve and enhance effective learning of learners [15].

We use literature, academic paper statistics, and content analysis methods to systematically sort out and summarize the various literature related to educational data mining that has been published in China and abroad and then conduct an overall study and analysis to objectively analyze and compare the current research status of data mining technology in the field of education at home and abroad [16, 17]. The research trend of educational data mining in the future is discussed. Through the above literature, we can see that data mining is more oriented to the improvement of algorithms of data mining technology and the mining of existing education management data to analyze the intrinsic relationship between their different attribute values, but there is relatively little research on the valuable information that affects the relationship between new students' arrival rate and education management decisions. The purpose of this paper is to analyze and study educational management data from this perspective using mining techniques [11, 18]. The development of an educational management information management system plays an important role in promoting the development of higher education informatization. Use data mining algorithms to establish models, test models, and application models to help educators understand the characteristics of education management as much as possible, analyze the factors that affect teaching quality, and maximize the value of data related to education management information as much as possible [19]. Give full play to the due role of educational data mining [20].

3. Research on the Decision Support System for Education Management Based on Data Mining

3.1. Data Mining Education Model Construction.
Reasonable goals in the decision-making process are a prerequisite for rational decision-making. The formation of the decision goal, the size of the goal, and the decision maker's understanding of the goal will affect the smooth implementation of the decision [21, 22]. In the process of establishing the goal, the nature, structure, and crux of the problem to be solved must be analyzed clearly before a reasonable decision target can be determined in a targeted manner. The decision goal must be very clear if the goal is too abstract or ambiguous, and ambiguous decision will be

difficult to carry out, and the degree of achievement of the decision goal is also difficult to measure. The conceptual representation of the educational management decision support process is shown in Figure 1. The education management decision-making process mainly includes four basic stages: target determination, design plan, implementation plan, and evaluation plan. Each stage of the educational management decision-making process has a close relationship with the decision-making environment.

After the decision objectives are determined, a variety of possible options are further designed for decision-makers based on feasibility studies. The proposed feasible options require a combination of overall exhaustiveness and mutual exclusivity to avoid bias in the option selection process. Overall exhaustiveness means that the various options developed should include all possible options found. Mutual exclusivity means that only one option can be chosen among different options [23]. These mutually exclusive options should all be easy for decision makers to compare and choose. The analysis, evaluation, and comparison of options mean that the indicators of all options are compared including the technical, economic, and social environmental conditions, factors, and potential problems, and dynamic financial indicators related to the decision objective are compared and evaluated. The possible constraints and potential problems of each option, as well as preventive and emergency measures, are also compared and evaluated. The proper execution of the user in the data mining process can accelerate the execution of the data mining process. Providing an interactive interface provides users with easy operation, and the interactive interface conveys the generated results to the user in a timely and convenient manner, and the generated results can be diverse.

The task of data integration is to merge data from multiple data sources into a unified data store for data mining. Data sources may include multiple databases, data cubes, and data files. The main considerations of data integration are attribute matching, data redundancy, and the detection and handling of value conflicts [24]. For attribute matching, attribute merging is performed by examining the field meaning of each attribute, data type, and so forth. Redundant data is considered at two levels. For attribute-to-attribute redundancy, the meaning of attribute entities is mainly examined, and attributes are streamlined by retaining fine-grained metrics of attribute values and ignoring coarse-grained metrics. This is because coarse-grained attribute data can come from the transformation of fine-grained data. Some redundant relationships between attributes are difficult to detect and can be detected by detecting how much an attribute can contain an attribute, such as by calculating the Pearson coefficient, also called the correlation coefficient of two attributes; for example, the correlation coefficient $R(X, Y)$ given two attributes X and Y is shown in equation (1), where N is the number of tuples, x_k and y_k are the values of X and Y in tuple k, respectively, $\bar{X}$ and $\bar{Y}$ are the means of X and Y, and $\beta(X)$ and $\beta(Y)$ are the standard deviations of X and Y. The ID3 algorithm is used to establish a decision tree model based on the attributes of the database samples. The algorithm requires more logarithmic
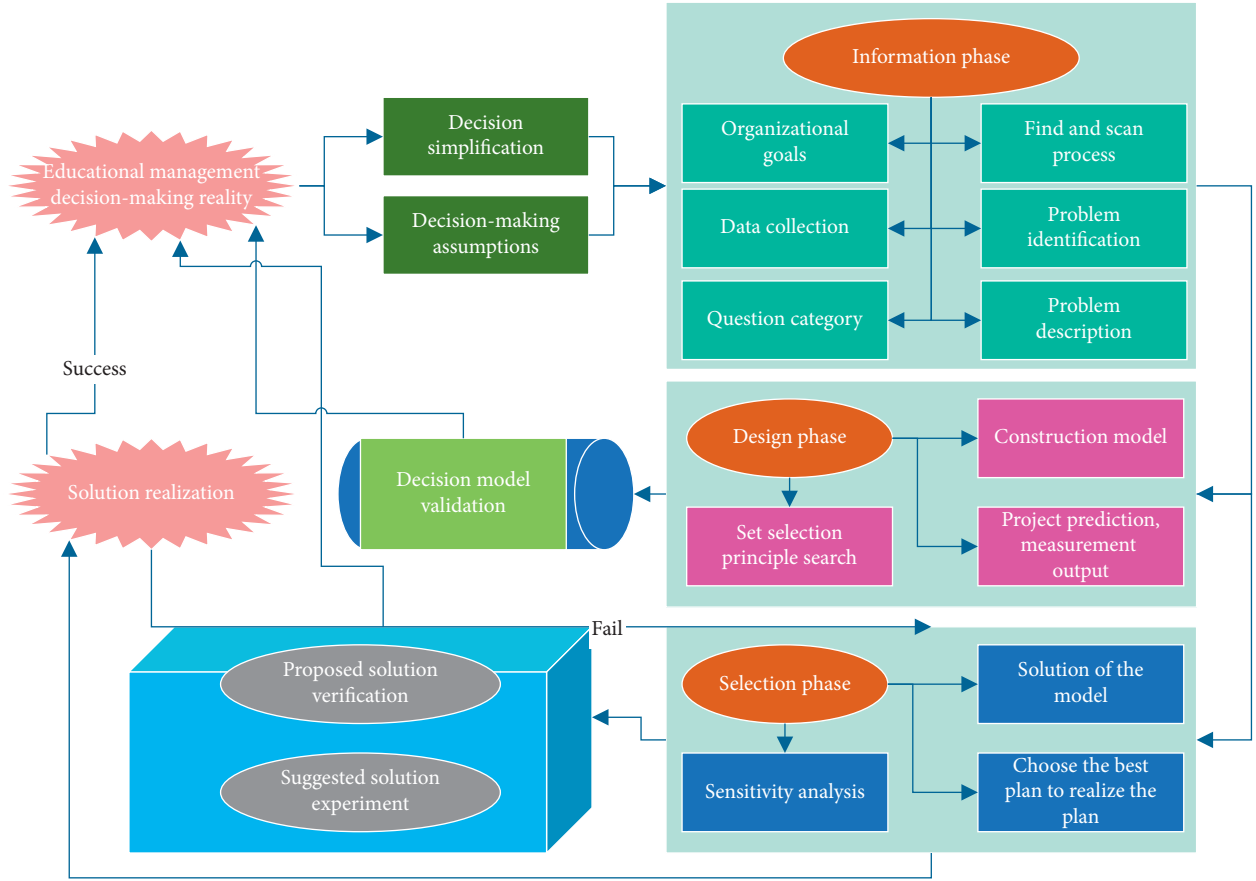
Figure 1: Educational management decision support process.

operations, which results in time complexity and low operating efficiency. The same formula needs to be calculated repeatedly, so the algorithm still has room for improvement. The algorithm needs to be optimized.

$$R(X,Y) = \frac{\sum_{k=1}^{N}\left(x_k - \overrightarrow{X}\right) * \left(y_k - \overrightarrow{Y}\right)}{N * \beta(X) * \beta(Y)}. \tag{1}$$

Facts are numerical measures, and multidimensional data models are organized around a fact, represented by a fact table that includes the name of the fact and the measure of the fact, as well as the code of each relevant dimension table [25]. A dimension is a pivot view or entity about which you want to keep records, and each dimension has a dimension table corresponding to it for data mining and decision analysis, which are Faculty Fact Table, Research Awards Fact Table, Research Results Fact Table, Research Funding Fact Table, Research Projects Fact Table, Talent Development Fact Table, Hardware Conditions Fact Table, and Subject Books Fact Table. Each fact table is associated with multiple dimension tables, such as discipline level, time, and unit level. The specific dimension tables are determined by the fact tables as shown in Table 1.

After obtaining the discrete dataset, we first train to obtain a complete decision tree using the ID3 algorithm, which works as follows: let data D be partitioned into a training set of class-labeled tuples, assuming that the class-labeled attributes have N different values, and N different classes H_i are defined ($i = 1, \ldots, N$). Let H_{iX} be the set of tuples of class H_i in X, and let $|X|$ and $|H_{iX}|$ be the numbers of tuples in X and H_{iX}, as in equation (2), where $F(X)$ is the average expected amount of information needed to classify the tuples in X; $|X_i| * F(X_i)$ represents the weights of the i-th division; H_i is the probability that any tuple in X belongs to class H_i and is estimated by $|X_i|/|X|$.

$$\begin{cases} F(X) = \sum_{i=1}^{N} H_{ix} * \ln(X_i), \\ \\ \overrightarrow{F(X)} = \sum_{i=1}^{N} \frac{|X_i| * F((X_i))}{|X|}. \end{cases} \tag{2}$$

The Bayesian posterior theory is used to test each branch of the decision tree, and the branches that are not considered to have sufficient generalization ability or are unreliable are removed from the tree, resulting in a more compact tree [26, 27]. The testing process is targeted at the knowledge of each rule translated from the decision tree, where each classification rule is obtained by searching from the root node of the decision tree top-down to a leaf node, and each classification rule consists of a conditional attribute tuple X and a conclusion classification label H_x. For the test of each classification rule, we define two kinds of validation.

TABLE 1: Fact and dimension tables for data mining data warehouse.

Serial number	Theme	Fact table	Dimension table 1	Dimension table 2	Dimension table 3
1	Scientific research awards	Scientific research award fact sheet	Unit dimension	Subject dimension	Time dimension
2	Scientific research results	Fact sheet of scientific research achievements	Subject dimension	Award type dimension	Unit dimension
3	Research funding	Research funding fact sheet	Award type dimension	Time dimension	Subject dimension
4	Research project	Research project fact sheet	Time dimension	Unit dimension	Subject dimension
5	Talent development	Talent training fact sheet	Unit dimension	Subject dimension	Award type dimension
6	Hardware condition	Hardware condition fact table	Subject dimension	Award type dimension	Time dimension

Adequacy verification:

$$G(Hi) = \sum_{i=1}^{N} \frac{|H_{ix}|}{|X|}. \tag{3}$$

Necessity validation:

$$\begin{cases} G(H_x|Y) = \frac{|H_{x,X}|}{|X|}, \\[2mm] G(H_x) = \frac{[H_x]}{|X|}, \end{cases} \tag{4}$$

We need knowledge with clear interpretability yet do not want the classification and prediction process to be too demanding, so to facilitate the simplification of the processing problem, we draw on fuzzy theory and perform some fuzzy optimization of the decision rules in classification and prediction. The affiliation function for a certain attribute value x belonging to a certain interval $(G_1, G_2]$ is established as

$$f(x) = \begin{cases} 1 - \dfrac{G_1 - x}{G_2 - G_1}, & x < G_1, \\[3mm] 1, x = G_1, x = G_2, & f(x) \subseteq [0,1], \\[3mm] 1 - \dfrac{x - G_2}{G_2 - G_1}, & x > G_2. \end{cases} \tag{5}$$

A minimum acceptable value of affiliation u is specified, which can also be determined from a large number of experiments. In this paper, the value is set to $u = 0.9$ to simplify the determination process.

$$f(x) = \max_{i \subseteq [1,N]} (G_1 - (1-u)(G_2 - G_1), G_2 + (1-u)(G_2 - G_1)). \tag{6}$$

The complexity of the educational system makes educational decision-making more difficult; an effective and rational educational decision support system will greatly contribute to the efficiency and rationality of educational decision-making. In this paper, we combine the characteristics of the educational decision-making process, the characteristics of the educational system itself, and the structural characteristics of the new generation decision support system to propose a decision support system model of educational management for the actual situation in the field of education and teaching.

3.2. Education Management Decision Support System Design Implementation. The data mining visualization system for education management information proposed in this study aims to provide decision support for education management information management in colleges and universities. The design goal of the system is to perform data mining on a large amount of education management data accumulated in the education management database over the years, and education management staff need to perform basic data analysis and data management on candidate data information through the online national college education management system. Finally, determine whether to admit or not and announce the admission results to the admitted candidates through the education management consulting system. The admitted candidates can also inquire about the situation of colleges and universities, previous years' education management admission score line, professional education management plan, application guidance, admission result, and issuance of admission notice through the education management information service system, while the relevant education management information managers of colleges and universities can also inquire about the historical information of education management and also use the data mining system to inquire about the application and admission and reporting situation through the system. Online analysis and visualization can be used to grasp education management information and make timely education management decisions.

The education management administration subsystem completes the function of maintaining information for education management records. All operations of the system require general records to be kept, and the general records are partially handled by the logs that participate in the shared section. The system starts when the user has successfully logged in and obtained the code that can control the scope of the system and expects to perform education management

arrangement management operations. Education management arrangement includes 6 common functions of form presentation, adding, changing, deleting, finding, and detailing of education management teaching maintenance data. The administrator selects the hyperlink of education management and clicks on the specific list of each education management arrangement, and, with the help of clicking the function button given in the upper right corner, the system is commanded to complete the addition, deletion, and change of the list of activity plans.

A large part of the data application uses the report query mode for data service, and so does the education management analysis module. The flow of the report query is briefly explained below, and its flow chart is shown in Figure 2. The first step of the report query is to log in to the system, at which time the system permissions are verified and only the functions and pages that the users are entitled to access are displayed; then the users enter the specific functional report interface according to their needs; the report interface functions are built based on various data analysis methods such as trend analysis, comparative analysis, cross-analysis, and retention analysis to establish the data model for query, and then the page is built based on the data model at the front end and the data organization at the back end. In the previous section, the detailed design and implementation of data storage and processing also mentioned that the data will be extracted from the dimension fields and indicator fields in the ADS layer, and now the query module can display the user-selectable dimensions based on the dimension field table. The user selects the dimension, the dimension value, and the data to send a query request, and the server calls the corresponding controller interface based on the requested URL and parameters to complete the query of the model data and return it to the Web client for display to the user.

The education management system designed and implemented in this paper uses a decision tree classification algorithm to generate a classification rule. One of the important tasks of the university's education management is to provide career guidance for students, so the education management system can use the information about students' schools to dig out their career direction. Therefore, the classification rules provided by the education management system for career guidance are as follows: Firstly, the students' school transcripts and comprehensive assessment are used as the input of the classification rules, and the courses taken by the students are classified as theoretical and practical courses; the spare time activities are classified as athletic and linguistic, and the students' roles in the activities are classified as leadership and nonleadership; based on this input information, the industry in which the students are employed is predicted. Finally, the employment information of previous students is used as an evaluation criterion. For example, grade analysis in university education management can obtain information about the knowledge points of students' grades in terms of question types and subject knowledge points of students' grades. It is very interesting to extract the important knowledge from the data warehouse of students' performance in higher education because there is a lot of information that can help to improve teaching management. The first step is to analyze the question types and the degree of correlation between question types, scores, knowledge points, and scores to summarize the strategies to train students.

4. Results and Analysis

4.1. Data Mining Education Model Analysis. To test the effect of the improved algorithm, the objective function can be scaled up and processed, and the number of misclassified entries can be calculated directly when calculating the error rate. The improved algorithm has an obvious effect on datasets with large data volume and complex data types, so the comparison of the convergence effect on the Vote dataset is listed, as shown in Figure 3, for the convergence curves of the algorithm before and after the improved algorithm is carried out for 40 experiments and taken as the mean value. Through the experimental comparison, for the dataset with large data volume and complex data type, multiple experiments are conducted, and, within 500 iterations, the improved algorithm converges significantly faster than the unimproved one, is less likely to fall into the local optimal solution, and can converge to the global optimal value early.

Figure 4 shows the error rate of classification. When the proposed algorithm is used for classification, the error rates of all the datasets used are generally lower than those of NBC and NBC-W. If there is a dataset with a high classification error rate on the test data, it may be that the parameters set are not suitable, the training samples are not set reasonably well, or the data satisfy the assumption of conditional class independence. In general, NBC-IBA can classify accurately and reach the desired accuracy when dealing with most of the datasets. In terms of the running time of the algorithm, since the NBC-IBA algorithm needs to iteratively search for weights, the training time of the classifier is a little slower than that of NBC and NBC-W. After the training, not only is the speed comparable to that of NBC but also the accuracy rate is substantially improved in the subsequent classification process. In this chapter, the normalization criterion is applied when converting positions to weights, and the globally optimal positions are not simply used as weights.

Figure 5 shows the comparison of the time run using the 3000 datasets collected in the object, and after the above analysis is applied to the actual one, from the perspective of the overall operation process, the advantages of optimization are very obvious, which not only reduces the number of calculation, but also calculates the order of all nodes, reducing the time complexity of operation. The optimized decision algorithm converts the original logarithmic operation into a simple four-rule operation during the calculation of the information gain, say application formula, which greatly reduces the average and overall computational effort. Therefore, the improved decision algorithm reduces the efficiency without changing the overall effect of the original algorithm and can completely replace the old algorithm for application to the system.
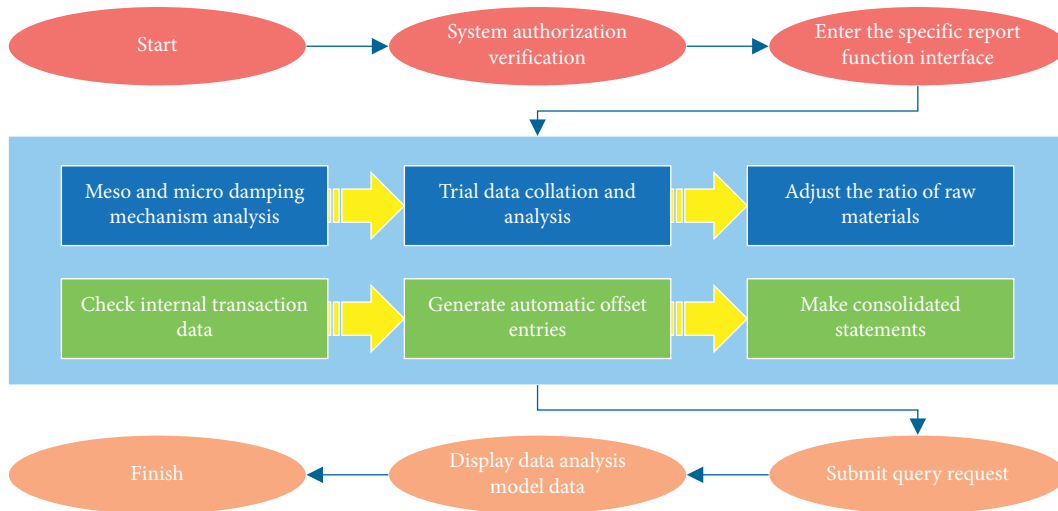
FIGURE 2: Flow chart of report query.
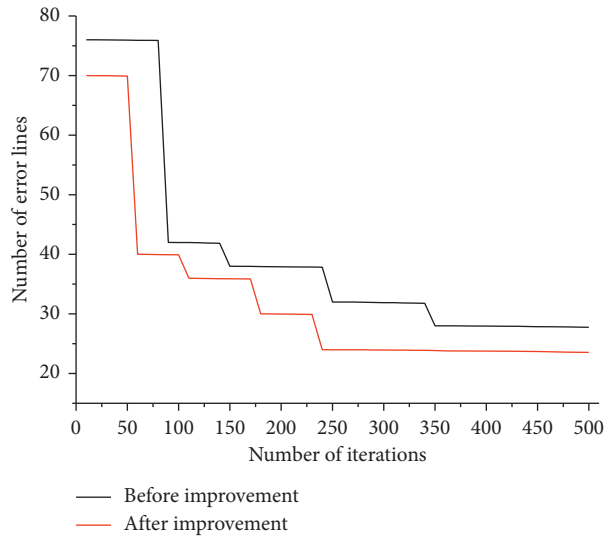
FIGURE 3: Convergence curves of the number of misclassified rows in the Vote dataset.

4.2. Analysis of Educational Management Decision Support System. Figure 6 shows the comprehensive analysis, which shows the basic information of 2017–2020, the students' reports of different majors, different places of origin, and different subjects, and the final educational decision can be made by combining the educational data of four years as follows. The system analysis module shows that, except for economic management and foreign languages, which have reached the education standard in the last three years, all other key majors have not been able to reach the basic number of education standards, so it is necessary to adjust the education number of majors in this area. The specialties of computer science, environmental engineering, and other science and technology majors can be adjusted appropriately according to the actual situation to make the specialties more rational and to get more students enrolled and admitted in this way. From the above conclusions, we can see that many factors are affecting the enrolment rate, and then we can use

this as a guiding basis in our future educational work to help the college decision-makers to plan the opening of courses and the adjustment of the setting of majors, as well as to give some scientific guidance to the work of discipline management departments as well.

Figure 7 further verifies that there is a significant difference in overall assessment scores between the preceding teacher-training category and the conductive category and between male and female students, while it can be known that teacher-training category is higher than the conductive categpry, and female students are higher than male students. In terms of comprehensive assessment, The students' comprehensive assessment in the academic year is linearly related to their intelligence, and ability has a significant impact on comprehensive assessment. There is a significant difference between male and female students, with an average score of 69.96 for female students, which is higher than the average score of 60.10 for male students; there is a
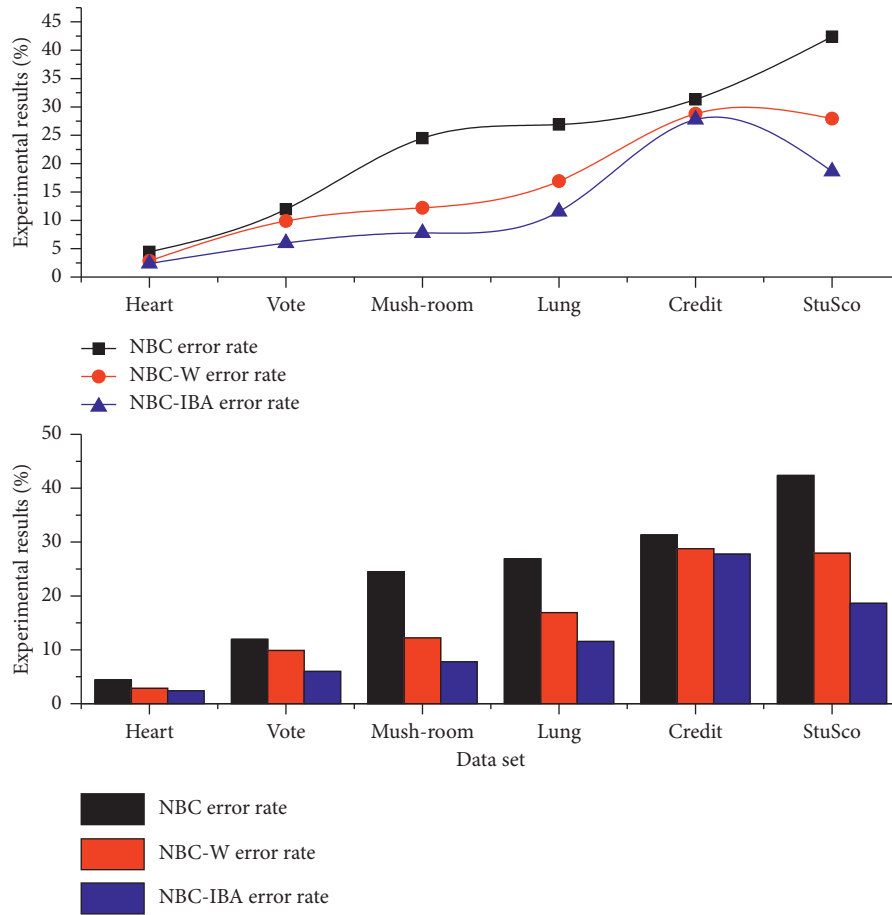
Figure 4: Comparison of classification error rates of three algorithms for six datasets.
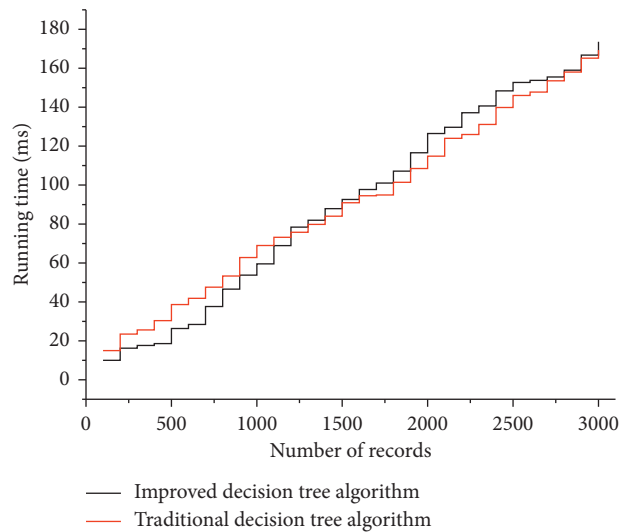
Figure 5: Running time.

significant difference between teacher training and knowledge, with an average score of 76.33 for teacher-training students, which is significantly higher than the average score of 55.28 for knowledge students. There is a significant

relationship between students' employment and their overall assessment and also with gender.

The statistical analysis results of the whole school are used as an example of the display, as shown in Figure 8,
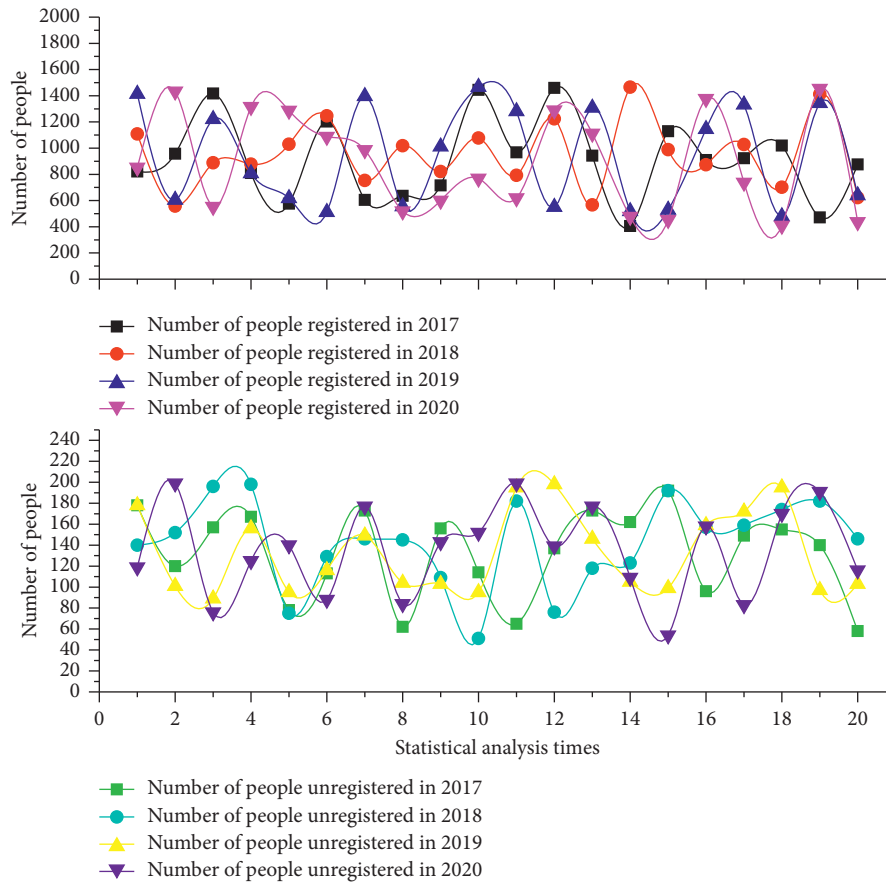
Figure 6: Comprehensive analysis.
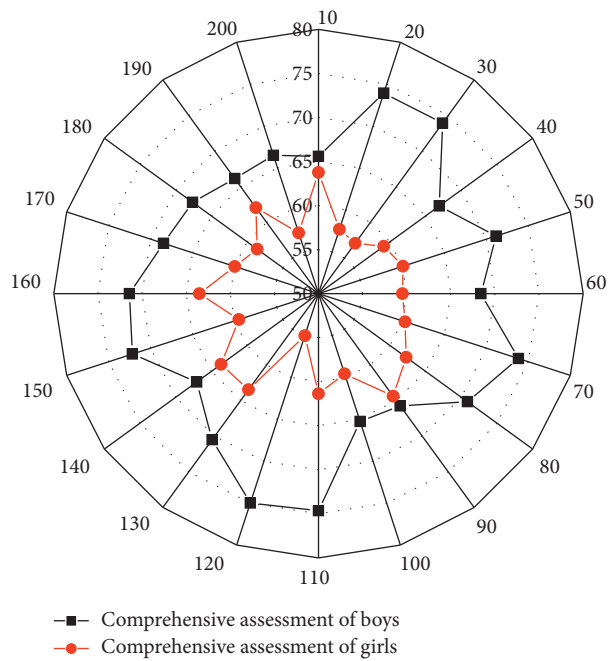
Figure 7: Correlation chart of comprehensive assessment scores.
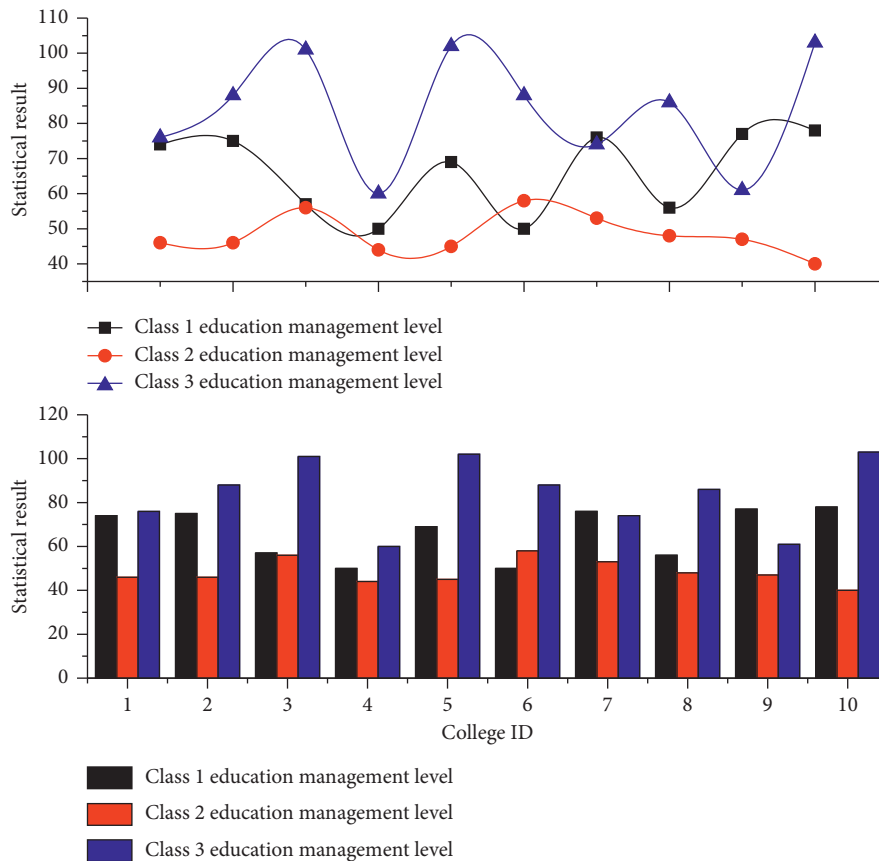
FIGURE 8: Statistical chart of educational management results by faculty.

which shows the statistical interface of the bar graph of education management results. The following statistical charts can be saved online. Figure 8 takes the statistical graph of education management of each college of the whole university as an example and shows the specific distribution of 3 types of education management levels of each college according to the college as a division unit, from which it can be seen which college is worthy of attention, which college is worthy of praise, and so on; the majors of a certain college, classes of a certain major, students of a certain class, and so forth are applicable to assist decision-makers in making decisions.

From the results of the above analysis, it is clear that there is a gap between male and female majors in educational technology in numerous aspects. To avoid this current gap phenomenon, the director of the program and the dean of instruction need to actively study and take favorable measures to guide male students to be competitive in academic life and other aspects to catch up with female students and narrow the gap. When setting the weighting coefficients for the comprehensive assessment of the student's academic year, the developers should focus on the coefficients for intellectual achievement and the criteria for ability bonus points. This is because these two parameters play a crucial role in the fairness and reasonableness of the comprehensive assessment of the student academic year. There is a need to continue to maintain the high weighting ratio of intellectual achievement in the

comprehensive assessment. This will promote the all-around development of students' moral and intellectual and physical abilities while adhering to the principle that "students' primary responsibility is to learn" and giving full play to the effective means of reflecting students' learning in their intellectual performance. Professional teachers and administrators should gradually guide students to make efforts to study and stabilize their English grades from the time they enter the university, to achieve a relatively smooth passage of the fourth and sixth grades, avoiding the clinical learning style and the need to explore effective teaching methods to improve the English grades of male students. Although the results of comprehensive evaluation have an important impact on students' employment, students' single ability or special skills have a great advantage in employment. At the same time, in the employment process, the employment opportunities for men are higher than those for women, this through interviews and analysis of the employment situation in recent years, and the reason is mainly generated by the employment units in the treatment of gender differences between men and women and the point of conformity. Various activities are vigorously carried out to increase the opportunities for students to demonstrate their abilities and encourage them to take various certificate examinations to expand employment opportunities. Because of the current employment situation, female students should choose suitable majors according to gender characteristics when they are employed.

5. Conclusion

With the transformation of knowledge-based economy, the education industry is developing rapidly in this context, and modernization of education management is an important basis for promoting the sustainable development of the education industry, and an important symbol of modernization of education management is the construction of information system, which is an inevitable choice for the development of education modernization. In recent years, urbanization construction has been making breakthroughs, the population in need of education is relatively mobile, and the school layout has to be adjusted continuously as a result, which undoubtedly brings about serious problems to the education authorities, so it is extremely crucial to build an education management decision support system based on data mining. In this paper, the significance and background of the educational decision support management system are explained, and the user requirements, design scheme, and functional implementation are systematically described. The overall architecture and functions of the system are also analyzed, as well as the related system development environment and the required database design. The key technologies and difficulties involved in the development of the system are also discussed in detail. Finally, through testing the system, we found that the system can meet the user's requirements and can complete the basic education management work and has certain decision support functions. Of course, due to my limited ability, some functions of the system are not perfect, and the development of the system is still in the initial stage, so problems and shortcomings are inevitable, which are also the areas that need to be improved in the future. It is necessary to carry out research on the construction and application of various prediction models and monitoring models for the system prediction and dynamic monitoring mentioned in the discussion of the role of educational decision support system for evidence integration and benefit integration. These research skills will be a further extension and deepening of this research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest in this paper.

Acknowledgments

References

[1] T. S. Kumar, "Data mining based marketing decision support system using hybrid machine learning algorithm," *Journal of Artificial Intelligence*, vol. 2, no. 3, pp. 185–193, 2020.

[2] E. Sugiyarti, K. A. Jasmi, B. Basiron, M. Huda, K. Shankar, and A. Maseleno, "Decision support system of scholarship grantee selection using data mining," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 15, pp. 2239–2249, 2018.

[3] A. K. M. Masum, L. S. Beh, M. A. K. Azad, and K. Hoque, "Intelligent human resource information system (i-HRIS): a holistic decision support framework for HR excellence," *International Arab Journal of Information Technology*, vol. 15, no. 1, pp. 121–130, 2018.

[4] M. W. L. Moreira, J. J. P. C. Rodrigues, V. Korotaev, J. Al-Muhtadi, and N. Kumar, "A comprehensive review on smart decision support systems for health care," *IEEE Systems Journal*, vol. 13, no. 3, pp. 3536–3545, 2019.

[5] B. Bakhshinategh, O. R. Zaiane, S. ElAtia, and D. Ipperciel, "Educational data mining applications and tasks: a survey of the last 10 years," *Education and Information Technologies*, vol. 23, no. 1, pp. 537–553, 2018.

[6] D. Dellermann, N. Lipusch, P. Ebel, and J. M. Leimeister, "Design principles for a hybrid intelligence decision support system for business model validation," *Electronic Markets*, vol. 29, no. 3, pp. 423–441, 2019.

[7] S. Sremac, E. K. Zavadskas, B. Matić, M. Kopić, and Ž. Stević, "Neuro-fuzzy inference systems approach to decision support system for economic order quantity," *Economic Research-Ekonomska Istraživanja*, vol. 32, no. 1, pp. 1114–1137, 2019.

[8] P. Elkin, D. Schlegel, M. Anderson, J. Komm, G. Ficheur, and L. Bisson, "Artificial intelligence: bayesian versus heuristic method for diagnostic decision support," *Applied Clinical Informatics*, vol. 9, no. 2, pp. 432–439, 2018.

[9] M. W. Rodrigues, S. Isotani, and L. E. Zárate, "Educational data mining: a review of evaluation process in the e-learning," *Telematics and Informatics*, vol. 35, no. 6, pp. 1701–1717, 2018.

[10] R. Ghorbani and R. Ghousi, "Predictive data mining approaches in medical diagnosis: a review of some diseases prediction," *International Journal of Data and Network Science*, vol. 3, no. 2, pp. 47–70, 2019.

[11] A. Farouk and D. Zhen, "Big data analysis techniques for intelligent systems," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 3, pp. 3067–3071, 2019.

[12] J. Du, H. Jing, K.-K. R. Choo, V. Sugumaran, and D. Castro-Lacouture, "An ontology and multi-agent based decision support framework for prefabricated component supply chain," *Information Systems Frontiers*, vol. 22, no. 6, pp. 1467–1485, 2020.

[13] J. Shen, L. Tan, L. Tan, and Y. Wang, "Design and implementation of international agricultural and biological engineering expert management system based on WEB mode," *International Journal of Agricultural and Biological Engineering*, vol. 13, no. 6, pp. 195–200, 2020.

[14] S. Lee, Y. Hyun, and M. J. Lee, "Groundwater potential mapping using data mining models of big data analysis in Goyang-si, South Korea," *Sustainability*, vol. 11, no. 6, pp. 16–78, 2019.

[15] C. Zhang and Y. Chen, "A review of research relevant to the emerging industry trends: industry 4.0, IoT, blockchain, and business analytics," *Journal of Industrial Integration and Management*, vol. 5, no. 1, pp. 165–180, 2020.

[16] J. R. Saura, "Using data sciences in digital marketing: framework, methods, and performance metrics," *Journal of Innovation & Knowledge*, vol. 6, no. 2, pp. 92–102, 2021.

[17] A. Prathik, J. Anuradha, and K. Uma, "Survey on spatial data mining, challenges and its applications," *Journal of*

Computational and Theoretical Nanoscience, vol. 15, no. 9-10, pp. 2769–2776, 2018.

[18] T. Van Nguyen, L. Zhou, A. Y. L. Chong, B. Li, and X. Pu, "Predicting customer demand for remanufactured products: a data-mining approach," *European Journal of Operational Research*, vol. 281, no. 3, pp. 543–558, 2020.

[19] Y. Nieto, V. García-Díaz, C. Montenegro, and R. G. Crespo, "Supporting academic decision making at higher educational institutions using machine learning-based algorithms," *Soft Computing*, vol. 23, no. 12, pp. 4145–4153, 2019.

[20] D. Solomon, S. Patil, and P. Agrawal, "Predicting performance and potential difficulties of university student using classification: survey paper," *Math*, vol. 118, no. 18, pp. 2703–2707, 2018.

[21] M. I. Habibie, R. Noguchi, M. Shusuke, and T. Ahamed, "Land suitability analysis for maize production in Indonesia using satellite remote sensing and GIS-based multicriteria decision support system," *GeoJournal*, vol. 86, no. 2, pp. 777–807, 2021.

[22] F. De Paola, M. Giugni, F. Pugliese, and P. Romano, "Optimal design of LIDs in urban stormwater systems using a harmony-search decision support system," *Water Resources Management*, vol. 32, no. 15, pp. 4933–4951, 2018.

[23] Q. D. Buchlak, N. Esmaili, J.-C. Leveque et al., "Machine learning applications to clinical decision support in neurosurgery: an artificial intelligence augmented systematic review," *Neurosurgical Review*, vol. 43, no. 5, pp. 1235–1253, 2020.

[24] B. Hmoud and V. Laszlo, "Will artificial intelligence take over human resources recruitment and selection?" *Network Intelligence Studies*, vol. 7, no. 13, pp. 21–30, 2019.

[25] A. Winter, S. Stäubert, D. Ammon et al., "Smart medical information technology for healthcare (SMITH)," *Methods of Information in Medicine*, vol. 57, no. S1, pp. e92–e105, 2018.

[26] C. Fischer, Z. A. Pardos, R. S. Baker et al., "Mining big data in education: affordances and challenges," *Review of Research in Education*, vol. 44, no. 1, pp. 130–160, 2020.

[27] S. Bharara, S. Sabitha, and A. Bansal, "Application of learning analytics using clustering data mining for students' disposition analysis," *Education and Information Technologies*, vol. 23, no. 2, pp. 957–984, 2018.

Research Article

Cloud Computing Image Recognition System Assists the Construction of the Internet of Things Model of Administrative Management Event Parameters

Peikun Xie ,[1] Enchen Ma ,[2] and Zaihua Xu [1]

[1]*Nanjing Forest Police College, Nanjing, Jiangsu 210046, China*
[2]*Southwest Jiaotong University, Chengdu, Sichuan 610031, China*

Correspondence should be addressed to Zaihua Xu; 106109@nfpc.edu.cn

In order to successfully apply the Internet of Things and cloud computing to the administrative management of spatial structures and realize the systematization, digitization, and intelligence of administrative management, this article draws on research experience in related fields and considers the data characteristics and computing tasks of administrative management. The whole cycle of transmission, storage, postprocessing, and visualization is the main line of research, and a cloud computing-based spatial structure administrative management IoT system is constructed. First, by summarizing the application status of the Internet of Things, the general Internet of Things system is summarized into three levels, and combined with the specific work in the spatial structure administrative management, the overall framework of the spatial structure administrative management of the Internet of Things system is proposed, and the functional sublayers are carried out. Secondly, in response to the above problems, through the traditional image recognition system research and practical application investigation, in order to meet the user's requirements for the computing efficiency and recognition accuracy of the image recognition system, an image recognition system in the cloud computing environment is proposed. It proposes a fuzzy evaluation algorithm of health grade hierarchy analysis optimized for the index system and scoring system and a calculation method that uses time series to identify regular outliers. The optical image pixel-level fusion method and the infrared and visible image fusion method based on complementary information are proposed, and the image fusion software is developed. Finally, in order to enable the application layer to use cluster resources to efficiently and intelligently process massive monitoring data containing redundancy, heterogeneity, anomalies, and many other defects, according to the calculation process of each specific task of data preprocessing and postprocessing in the application layer, demonstrations are made one by one. After analysis, it is concluded that vertical storage of data blocks according to different sensor channels is the optimal strategy.

1. Introduction

With the increasing popularity of smart sensing and wireless networking technology in the field of structural administration management, large-span spatial structure administration projects involving major public safety have grown rapidly in recent years [1]. Facing the continuous accumulation of massive amounts of heterogeneous monitoring information, with cloud computing as technical support, the establishment of a spatial structure administrative management Internet of Things system to achieve

comprehensive perception, intelligent processing, and visual design of monitoring data has become an inevitable requirement for administrative management construction [2]. They study the organic combination of the Internet of Things, cloud computing, and spatial structure administrative management, establish a spatial structure administrative data management cloud platform, give full play to the advantages of cloud computing and the Internet of Things system, and provide technical support for the storage and postprocessing of monitoring data. It is conducive to the accumulation of research forces, reduces the analysis cost of

monitoring data, and improves the scale and comprehensive benefits of the administrative management system [3–5].

With the injection of powerful and flexible storage and computing resources of cloud computing, the collaboration of smart hardware and software will become closer. The Internet of Things system will gradually be different from the traditional collection and display system and become a highly information-based organic whole. With the continuous improvement of wireless sensor technology and intelligent control technology, the intelligent, automated, integrated, platform-based, and visualized administrative management of spatial structure has become an inevitable trend in the development of the industry [6]. The Hadoop platform is a parallel programming model and computing framework for processing massive data and is used for parallel computing on large-scale data sets. Grid computing is a technology that has gradually emerged and developed in recent years. It is currently widely used in many research fields such as distributed supercomputing, distributed instrument systems, data-intensive computing, and remote immersion. Using the supercomputing power of grid computing to solve the problem of image recognition is also one of the current research directions. By improving the image recognition algorithm to the nonparallel algorithm, the program is transplanted to the grid computing platform to run, and the distributed image recognition technology is realized [7]. It has a simple structure, can effectively support data-intensive applications, and is very suitable for the development of parallel digital image processing technology. The Internet of Things is an important part of the new generation of information technology and has become the third wave of the information industry after computers, the Internet, and mobile networks [8–10].

Cloud computing platforms and cloud computing network operators are at the core of the entire dynamic administrative management construction process. When an order comes, the corresponding historical data is processed through the cloud computing center, and the optimal node for the corresponding order is quickly selected for the formation of the Internet of Things. Dynamic administrative management responds to demand, so as to achieve the purpose of improving the overall efficiency of dynamic administrative management. The clustering algorithm is introduced at the transport layer to ensure high-efficiency transmission and effectively reduce data loss; at the application layer, data security levels are classified, and data of different security levels are protected to different degrees. In this strategy, the distribution of permissions must go through strict identity verification, which can avoid the security problems caused by illegal users obtaining illegal permissions in the traditional system. In addition, supporting secondary development is an inevitable requirement for the development of the Internet of Things. It is possible to classify the things in the Internet of Things to facilitate the design of users to meet their own needs. Choosing public cloud services as the development support, we establish a spatial structure administrative management cloud computing network and complete the development of the application layer of the Internet of Things system. The platform responds to various needs of users through internally encapsulated application services and realizes real-time reception, dynamic display, and comprehensive management of monitoring information, thereby establishing a complete spatial structure administrative management IoT system.

2. Related Work

In recent years, with the rapid development of communication networks and sensor components, automated and intelligent collection of monitoring data has been realized in structural administrative management technology, which also provides feasibility for real-time online analysis. In order to achieve real-time or even premastering of the state changes of structural performance, in addition to the need to improve the monitoring and collection subsystem to ensure accurate collection and transmission of information, it is also necessary to consider the interaction with users and develop a set of data mining, data display, and data storage. The monitoring data management system is integrated with other functions, so as to analyze and process the monitoring information online in real time and present the processing results to the monitoring users to assist the corresponding decision-making.

Kumari et al. [11] proposed a deep belief network model that can use an algorithm called greedy layer-by-layer training strategy for effective training, which quickly aroused people's enthusiasm for cloud computing network research. Subsequently, Sharif et al. [12] proposed that this training method is also applicable to other different types of deep neural networks and can systematically improve the generalization ability of the network model on the test samples. The convolutional cloud computing network is also a large category of cloud computing. Pasquier et al. [13] first introduced the application of the supervised backpropagation algorithm to this multilevel structure of the convolutional cloud computing network. In terms of speech recognition, Anuradha et al. [14] use the DNN-HMM architecture to replace the original GMM-HMM architecture and use the filter bank features of the original Mel domain to replace MFCC. It even dropped by half. In addition, cloud computing has also achieved high-profile results in pedestrian detection and image segmentation and has surpassed human levels in traffic sign classification. For the impulse cloud computing network, starting with the discovery of impulse neurons, Elhoseny et al. [15] proposed the first bionic impulse neuron model. Although its biological nature is very strong, the computational complexity is too high, so there is subsequent evolution of FHN, ML, IZH, IF, and other neuron models, thus greatly reducing the computational complexity. The learning rules for impulse cloud computing networks have gradually evolved from the original Hebb learning rules to STDP and IP mechanisms, thereby improving the learning efficiency of impulse cloud computing networks [16–18].

The classification and recognition accuracy of the image recognition system largely determines the quality of the entire system. Therefore, in image recognition technology, classification and recognition technology should be paid

enough attention to. Through the research and improvement of classification and recognition technology, the computational efficiency and recognition accuracy of image recognition can be improved, and the foundation for wider applications can be laid [19]. At present, the main classification techniques include decision tree induction, Bayesian classification, KNN classification, and artificial cloud computing network. Among these methods, KNN classification is a simple, effective, nonparametric method and has been widely used in text classification, pattern recognition, image and spatial classification, and other fields. The KNN algorithm is a lazy learning method, where the learning program will construct a model until the last moment before classifying the given test set. In classification, the computational cost of this learning method and the need for large storage cost were needed [20, 21].

3. System-Assisted Administrative Management Event Parameter IoT Model Construction Based on Cloud Computing Image Recognition

3.1. Cloud Computing Hierarchical Architecture. Various services of the cloud computing system exist in the network, users do not need complicated procedures, they only need to access system resources through the Internet, and the cloud computing system supports various access methods, allowing users to access various clients. The information transmission between the three levels of the Internet of Things system is not simple one-way transportation, but a two-way linkage intelligently controlled through a preset real-time feedback mechanism, which involves security technology, analysis and identification technology, service quality management, and cross-level public technology including network management [22]. Figure 1 is the cloud computing hierarchical architecture topology.

Cloud computing is a resource utilization mode, which can access configurable computer resource pools (such as networks, servers, storage, applications, and services) through the network in a convenient, friendly, and on-demand access manner. In this mode, we can quickly supply and provide services with minimal management costs. The main purpose of integrating physical resources is to integrate the services that these physical resources can provide and transform them into resource pools that provide users with different services, including computing resource pools, storage resource pools, network resource pools, and data resource pools.

$$L \times 1 + \frac{1}{L+d+N} \approx 1 - \frac{1}{L+d+N}. \tag{1}$$

The perception layer is the lowest layer in the three-tier system architecture of the Internet of Things, and it contains two application sublayers: data collection and short-distance communication. Its function is to perceive the behavior of the target thing through the sensor node, obtain the information data of the physical world, and then use the self-integrated communication protocol to organize the data exchange and transfer it to the upper layer.

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1m} \\ m_{21} & m_{22} & \cdots & m_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ m_{m1} & m_{m2} & \cdots & m_{mm} \end{bmatrix}. \tag{2}$$

The SOA building layer combines the services that cloud computing can provide to users and presents them to users through interfaces, which mainly include service interfaces, service registration, service search, service access, and service workflow. Cloud computing has good interfaces and can realize good interactions with various services.

$$l = \alpha \times \sum_{i=0}^{S^2} \sum_{j=0}^{B} \left[(x_i - \overline{x})^2 + (y_i - \overline{y})^2 + (x_j - \overline{x})^2 + (y_j - \overline{y})^2 \right],$$

$$L_{(\text{emg},k)} = \sum_{f(\text{emg},k) \in A_{(\text{emg},k)}} \left(f_{(\text{emg},n)} + C_{(\text{emg},k)} \right),$$

$$\min \sum_{i=1}^{N} a_i = \frac{1}{2} \sum_{i=1}^{N} \sum_{j}^{N} a_i \times y_j \times a_j \times y_i \times k\left(x_i^2, x_j^2\right). \tag{3}$$

As the middle layer of the Internet of Things system architecture, the network layer is mainly used to provide a long-distance communication medium for the two-way information transmission between the perception layer and the application layer on the premise of ensuring the security and integrity of the data. The development of the existing Internet long-distance communication technology has been relatively mature. As long as the data connection with the private network is completed through related protocols, the Internet can safely and in real time transport massive amounts of data in both directions.

3.2. Image Recognition Auxiliary Algorithm. From the actual operating point of view, image transformation is to find a suitable orthogonal transformation kernel for the original image. In essence, image transformation has a profound physical background. For example, a Fourier transform of the image reflects the frequency distribution of the function on the system spectrum. Generally, imaging systems only have a certain range of brightness, and the ratio of the maximum value to the minimum value of the brightness becomes the contrast. Using digital image processing technology and pattern recognition theories, combined with computer technology, we analyze various formed elements in urine sediment and find out reasonable segmentation methods and classification rules in accordance with digital image processing algorithms. The system that forms the image has limited brightness, so the contrast is not enough, so the human eye has a poor visual effect when viewing the image. The visual effect is improved through grayscale transformations such as linear transformation, piecewise linear transformation, and nonlinear transformation. Geometric normalization is also called position calibration. It will help correct the size difference and angular deviation caused by the imaging distance and
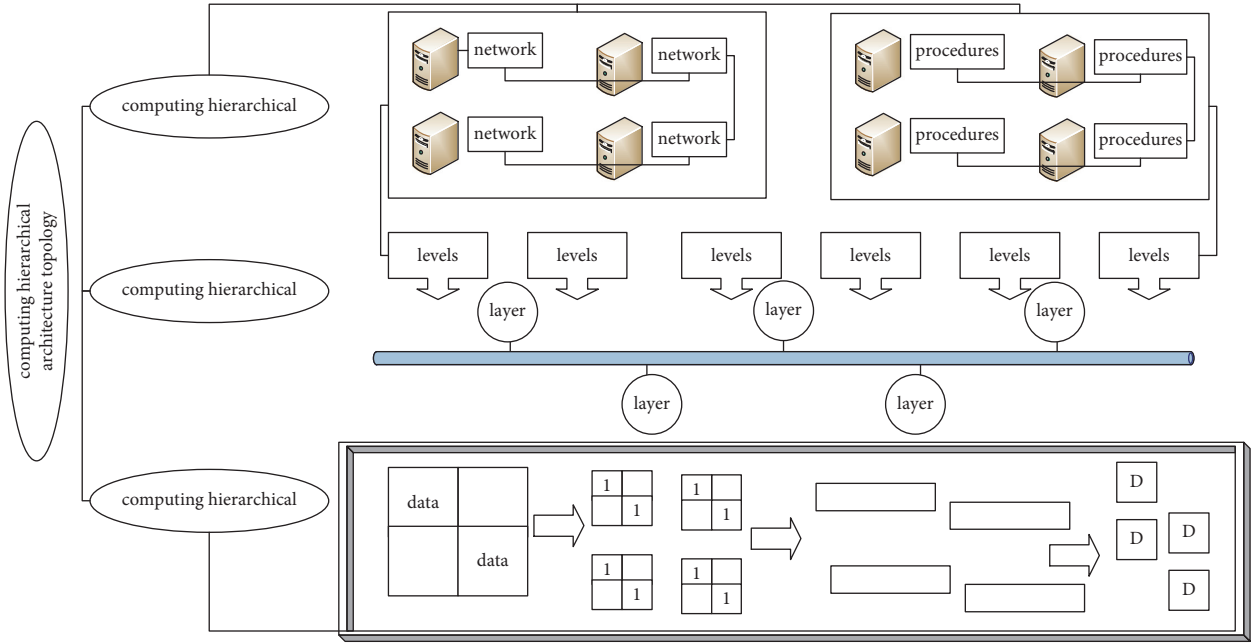
FIGURE 1: Cloud computing hierarchical architecture topology.

image pose changes. Gray normalization is used to compensate the image obtained under different illumination and light source directions, so as to reduce the change of the image signal caused solely by the illumination change. This step completes the task of finding the image area in the input image. After an image is given, we check whether there is an image in the image, and if so, we give the position and range of each image. Because image detection requires detecting the existence of images and determining their position from various scenes, it is also called image positioning (positioning an image in a scene is also called segmentation). If a static image is an input, each image will be tested. If the input is video, you need to obtain the image area in every frame in the entire sequence. In addition to detecting each frame as a static image, the common strategy is to detect only the first frame of the sequence and track the following frames with the detection results of the previous frame.

Figure 2 is a pie chart of image recognition input information. External input information is transmitted to the cloud computing network through the input layer. The neurons in the input layer transmit information to the hidden layer through a certain relationship. The hidden layer is responsible for information processing. The hidden layer can be set to a single implicit according to data processing needs. After the information is processed, the hidden layer passes the information to the output layer neurons through a certain relationship, and the output layer further processes the data and outputs the information processing results to the outside. The process of inputting external information from the input layer to outputting the output layer is a process of forward transmission and learning of information. If the actual output of the cloud computing network does not meet the expected output requirements, the network will propagate the error back to the output layer,
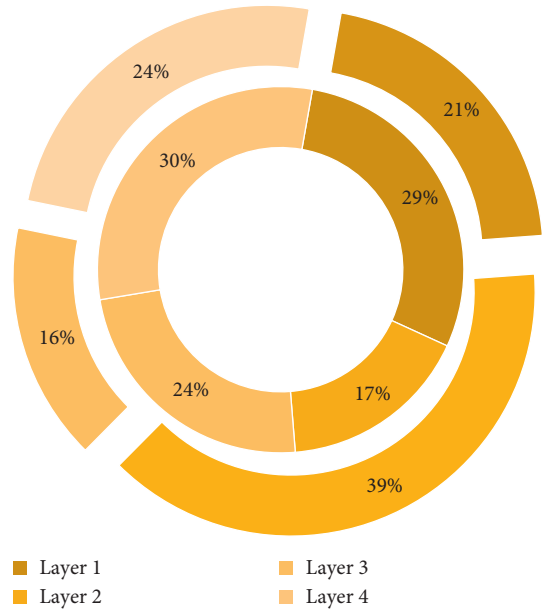
FIGURE 2: Fan graph of image recognition input information.

hidden layer, and input layer, layer by layer, and adjust the thresholds of each layer and the adjacent two layers according to the error gradient descent method. The above process is continuously repeated until the output error is reduced to the desired level or reaches the set step. This module is based on the foundation of the subspace image recognition module, provides the function of calculating and storing data, and is the guarantee of the system's operational efficiency. The distributed structure of cloud computing improves the computational efficiency of image recognition algorithms to use HBase to store the image and its feature values for classification services.

3.3. Parameter Allocation of Administrative Management Events. Users in the administrative area can obtain the services they need from a complete computer infrastructure through the Internet. IAAS is a business model that allocates hardware resources such as data centers and infrastructure to users through the Web. The advantage of infrastructure services is that users do not need to purchase high-quality equipment but only need to build an application system that meets their needs through Internet leasing. Compared with a single agent, MAS has many unique characteristics: because multiple agents can work in parallel, they can solve the problem very quickly; the communication between the agent and the agent does not need to use raw data but adopts a high-level communication language. This greatly reduces the communication traffic; the function of a single agent is relatively simple, and the agents in the MAS system can cooperate with each other and share tasks to solve some complex large-scale tasks when a problem occurs in a certain agent in the MAS system. Compared with the traditional model, for an enterprise and organization, cloud computing can put the data center in the cloud, and professional companies can provide them with different levels and types of information services. When the responsible task fails, MAS can quickly introduce other agents or replace them with other agents to complete the corresponding tasks, which greatly improves the reliability of the system. Distributed computing technology is also the core of cloud computing technology, which is mainly developed and utilized under the premise of factors such as availability, reliability, and economy. The representative applications of distributed data storage systems in cloud computing mainly include GFS (Google file system) and HDFS (Hadoop distributed file system).

Cloud computing is a comprehensive computing method that combines parallel computing, distributed computing, and grid computing. The development of Internet of Things technology must rely on technologies such as high-efficiency storage and high computing power. Figure 3 shows the parameter assignment of administrative management events. The Internet of Things combined with cloud computing technology collects and organizes data and information through smart devices such as wireless sensors and radio frequency identification and then transmits it to the cloud computing platform at the application layer to realize data sharing and exchange and complete the control and management of the entire system. By selecting an appropriate matching strategy, the image to be recognized can be matched and compared with the known image in the database, the correlation between them can be established, and the judgment decision made can be output. There are two kinds of recognition purposes and situations that need to be distinguished: one is to verify the image, that is, to confirm whether the person in the input image has an image in the database, which belongs to supervised recognition; the other is to recognize the image, namely, to confirm the identity of the person in the input image, which belongs to unsupervised recognition. Based on the summarized

functional requirements analysis, the spatial structure administrative management IoT system should consist of the following parts: a perception network responsible for the collection and control of project site monitoring information and file integration to achieve reliable data between the project site and the remote monitoring and control room transmission network channel according to the user's instructions to call resources to perform various data analyses, and at the same time responsible for the intelligent processing cluster for data storage to provide users with rich visual design and data application display terminals and to realize the user's manual connection with the underlying data through the data interface management terminal for identification and modification.

3.4. Weight Update of IoT Model. Hadoop under the Internet of Things model is a software framework that can perform distributed processing of large amounts of data. It implements Google's MapReduce programming model and framework. It can divide applications into many small work units and put these units in any cluster. In MapReduce, an application that is ready to be submitted for execution is called a "job," and the unit of work divided from a job that runs and each computing node is called a "task." In addition, the distributed file system (HDFS) provided by Hadoop is mainly responsible for data storage on each node and achieves high throughput data read and write. In terms of distributed storage and distributed computing, Hadoop uses a master/slave architecture. We run Hadoop through different background programs, which are composed of NameNode, DataNade, JobTracker, and TaskTracker. The NameNode and JobTracker are run on the Master node, and a DataNade and TaskTracker are deployed on each node so that the data processing program running on this node can directly process the data of the machine as much as possible. Other complex issues in parallel programming, such as load balancing, fault-tolerant processing, distributed storage, job scheduling, and network communication, are all handled by the MapReduce framework. The MapReduce processing data set has the following characteristics: the data set to be processed can be decomposed into many small data sets, and each small data set can be processed completely in parallel. Figure 4 shows the weight distribution of the image recognition model.

The application layer is the top layer in the Internet of Things architecture and the core layer in most Internet of Things systems. The application layer needs to intelligently process and comprehensively manage the collected massive data and provide user-oriented human-computer interaction and data visualization design. The application layer can be divided into an application support sublayer and an application service sublayer. The tasks to be processed by the application support sublayer include postprocessing applications such as data storage, mining, analysis, query, and search. On the basis of data processing, we realize cross-industry, cross-regional, and cross-system information
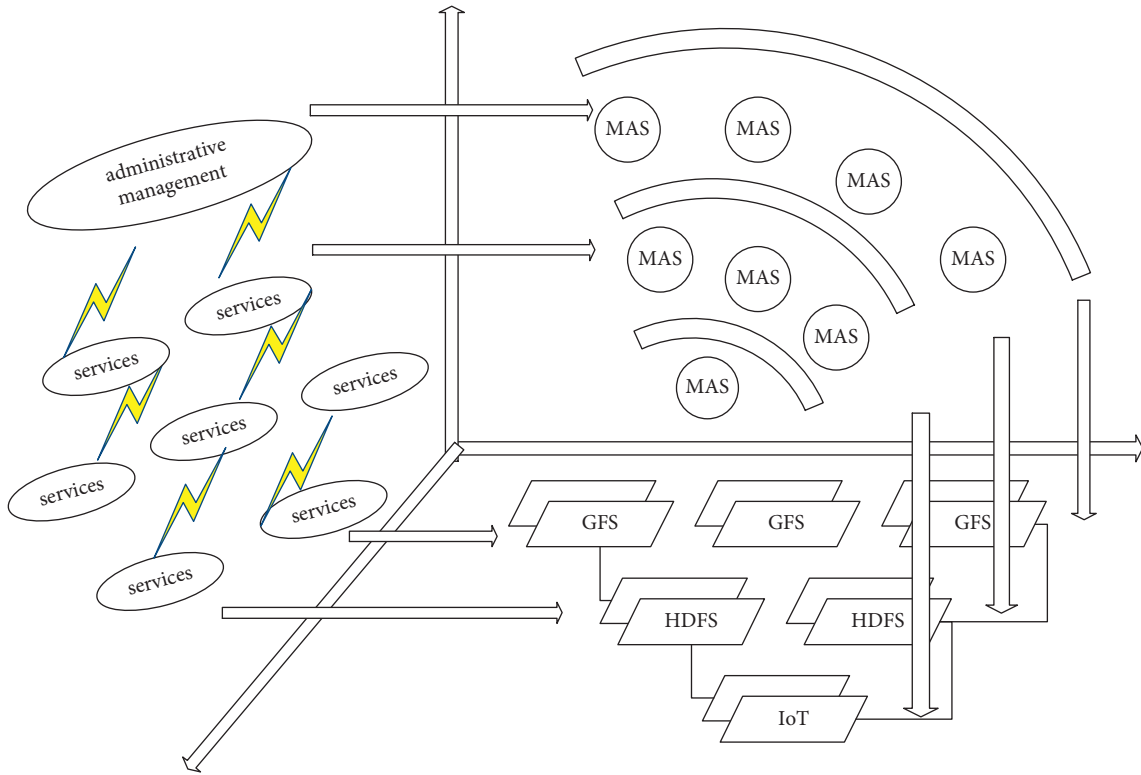
Figure 3: Parameter allocation of administrative management events.
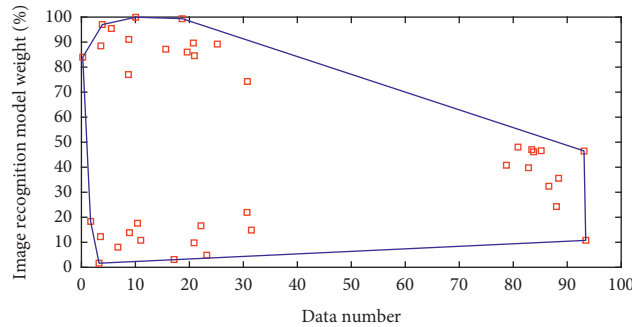
Figure 4: Weight distribution of image recognition model.

sharing, serving individuals, and organizations in different industries; the application service sublayer is a human-computer interaction interface set up for users to realize the interaction between users and applications order issuance and result feedback. Offline processing in an image recognition system refers to a series of preparations that the image recognition system needs to complete before accepting user queries to recognize images and pictures. This part of the work solves the problem of preprocessing and feature extraction of massive image data, mainly including image and picture preprocessing module, feature value and feature vector extraction module, and feature value access. The main feature of offline processing is that the real-time requirements are not high, but the image and picture index needs to be established in advance.

4. Application and Analysis of System-Assisted Administrative Management Event Parameter IoT Model Based on Cloud Computing Image Recognition

4.1. Cloud Computing Data Feature Extraction. The network structure of this experiment uses a four-layer deep belief network, the input layer is 784 neurons, the second and third layers as hidden layers each contain 500 neurons, and the last layer is the label layer and contains 10 neurons. We, respectively, represent the network weight between each layer. The Siegert neuron is used for network training, and the training method is the CD algorithm. The experiment content is divided into two parts. The first part is the

recognition mode, and the second part is the generation mode. The recognition mode is to give a gray image of handwritten digits to recognize numbers, and the generation mode is to give numbers and let the network generate handwritten digits in reverse. When the feature has a small displacement, the complex cell can ignore this effect, and the information obtained remains unchanged. This feature correspondingly achieves the pooling operation in the convolutional neural network. The grayscale image is reconstructed from handwritten digits. When testing the recognition speed of a single picture in the recognition mode, the LIF neuron is used, and the gray value of the picture needs to be normalized to a value between 0 and 1, and the pulse sequence is determined according to the proportion of each pixel value. In order to further enhance the consistency of the image mode, the statistical characteristics of the sample pictures can be normalized, considering the most basic statistics—the mean value and variance of the grayscale—and adjusting them to a given value.

Figure 5 is the feature vector curve of image recognition calculation data. When the image boundary is relatively smooth, the area density is smaller, and its density C is approximately equal to 1. When the shape of the image area deviates from the circle, the C value is smaller. When the edge of the image changes drastically, the perimeter of the image increases, but the area is relatively reduced, resulting in an increase in area density. Therefore, the area density C reflects the characteristics of the target topography to some extent. Here, we use edge detection to segment the subimages, because the canny operator has the characteristics of single edge response, accurate positioning, and no false edges. At the same time, it has the first and second derivatives of the image in the area edge extraction advantage. The specific steps are as follows: first, perform median filtering on the subimage extracted in the original image; then, perform image enhancement on the filtered subimage; secondly, use the canny edge detection operator to segment the enhanced subimage. Then, we use the morphological open operation and close operation processing and, finally, remove the small connected domains in the subimage.

4.2. Image Recognition IoT System Simulation. Handwritten image recognition is a classic experiment to verify the accuracy and efficiency of the network model. The data set used is the MNIST data set. The training data includes 60,000 handwritten digital grayscale images, and the test set includes 10,000 handwritten digital grayscale images. The image content is a number from 0 to 9, and the pixel size is $28 * 28$. Online processing refers to a series of tasks that the image recognition system needs to complete for the user to feed back the query results after the user enters the image picture or the person's identity information. This part of the work mainly includes the following: according to the image picture that the user needs to query or the person's identity information that needs to be queried, we find the corresponding person's identity information or image pictures from the HBase image database index file. The salient feature of online processing is that the real-time requirements are
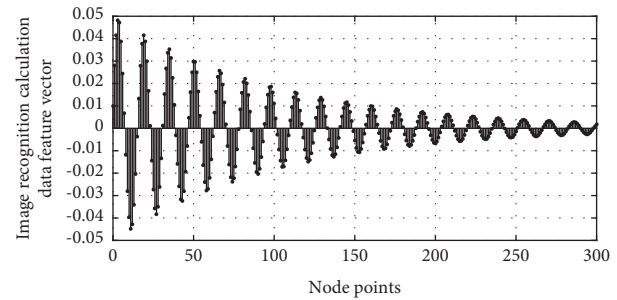
Figure 5: Image recognition calculation data feature vector curve.

very high, but all the image and picture data must be compared, so the calculation efficiency is relatively high. In this system, the video capture terminal and the front-end processing terminal are configured in the front-end monitoring, and the monitoring video is collected and obtained through the video capture terminal. The obtained video information can be directly transmitted to the cloud computing system through the transmission network or realize information processing through the front-end processing terminal. Generally, in order to ensure the robustness and usability of the system, multiple video capture terminals from different angles are generally set up in the monitoring area, and the coordination between the terminals is realized through the configuration of monitoring and management software.

In order to port the image recognition application to embedded devices, this section selects the Raspberry Pi 3rd Generation B-type and USB camera as the hardware device. The Raspberry Pi hardware uses the ARM Cortex-A53 chip as the processor, and the operating frequency is 1.2 GHz. Normally, when using a computer to process data, the continuous time signal will be discretized, and the given value is the data at a specific time interval. So, a more realistic assumption is that the sensor will give a measurement result every second. Although the Raspberry Pi is very small, it also has functions such as WIFI and Bluetooth. It can also communicate with other devices through GPIO pins. Figure 6 shows the accuracy of image recognition based on the Internet of Things. After adding IP, the recognition accuracy of the entire network has improved. It can be clearly seen that 94% is a demarcation point. When the initial discharge rate is greater than 700 Hz, the accuracy fluctuates on the line, but the final value is greater than 94%. In contrast, no IP mechanism was added. When the initial discharge rate was greater than 700 Hz, the recognition accuracy showed a fluctuating downward trend. It can be seen that the time consumed to recognize a single handwritten digital grayscale image is related to the initial discharge rate. After adding IP, when the initial discharge rate is greater than 200 Hz, the average time to recognize a single handwritten digital grayscale image is about 2.5 ms. Without adding IP, when the initial discharge rate is 1000 Hz, the recognition speed is about 5.8 ms. From this analysis, it can be seen that this IP mechanism has improved the recognition accuracy of the network, and the main impact lies in the improvement of the network recognition speed. Then, we want to know why the
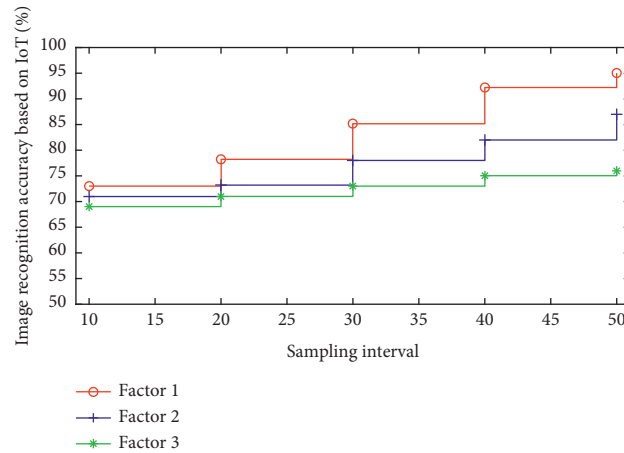
FIGURE 6: Image recognition accuracy based on the Internet of Things.

IP mechanism has such a big impact on the network recognition speed, and it can be seen that when the initial discharge rate is greater than 200 Hz, the broken line tends to be stable, with little fluctuation up and down.

4.3. Example Application and Analysis. In the cloud computing network, the ICP adopts a unit modular design and a universal platform that can control and connect a variety of different types of communication systems such as wired and wireless. As the core control device, ICP has dedicated interfaces for different communication methods, and each interface corresponds to a specific communication access device. The ICP interface supports an analog trunking interface, conventional intercom interface, shortwave radio interface, and GSM/CDMA mobile phone interface. Each interface independently completes the receiving and sending of audio signals, and at the same time, it can achieve complete control of the equipment for specific access devices. In order to improve the voice quality, all voice interfaces adopt a four-wire processing method to separate the sending and receiving signals to avoid mutual at the same time; in order to realize that each port can perform two-stage dialing, the external interface in the device has the independent DTMF signal detection and forwarding capabilities. Figure 7 shows the trend of image recognition transform and decomposing signals.

The process of using wavelet transform to locate the high-frequency components is as follows: decompose the image by wavelet transform, then set the low-frequency part to 0, keep the high-frequency part unchanged, and then perform inverse wavelet transform to complete the localization of the high-frequency components of the original image. In this article, the Mallat algorithm is used to decompose the image, by using Daubechies 9–7 wavelet to obtain 4 subband parts and then zeroing the low-frequency part to ensure that the high-frequency part remains unchanged; then, the reconstruction is performed, and the high-frequency part can be obtained. At the same time, we can select a suitable frequency value, which is between the defocus component, background, and component, and use this frequency value as a threshold for threshold processing.
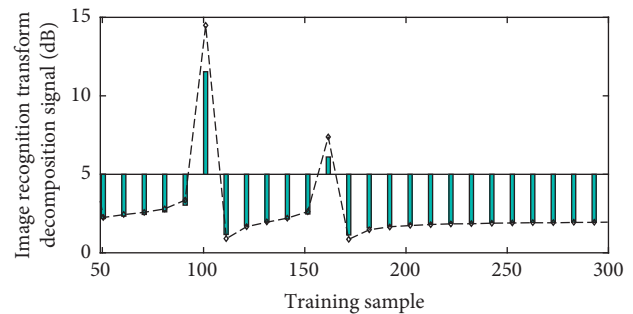
FIGURE 7: Image recognition transformation and decomposition signal trend.

The value above the threshold is 1; otherwise, it is 0; then the place with dense brightness is urine sediment. The position of the components is finally based on the empirical threshold of the image, and the corresponding binary image is obtained by simple discrimination.

This paper uses the VGG network pretrained on ImageNet as a model for extracting image visual features. The 1 million images provided by ImageNet make VGG can train more accurate features than manual design. In order to make the feature more versatile, this study uses the 4096-dimensional vector output by the layer in VGG-16 as the visual feature of the image. Figure 8 is a histogram of the frequency response of image recognition visual features. The input layer continuously receives the stimulation of the Poisson pulse sequence over time. When the neuron membrane voltage of the input layer exceeds the threshold, the neuron fires and transmits the pulse to the next layer of neurons. When the third layer of neurons fires, the final result is displayed in the last layer, and the last layer is a $2 * 5$ matrix, which means 0, respectively. The pooling process also has the function of adjusting the output size. In practical applications, for the convolutional neural network model, when the size of the input image is different, in order to obtain a uniform size output, it needs to be realized by pooling operation at this time. The 10 numbers to 9 are numbers, among which the white dots represent the firing neuron and the corresponding identified number. It can be seen that the
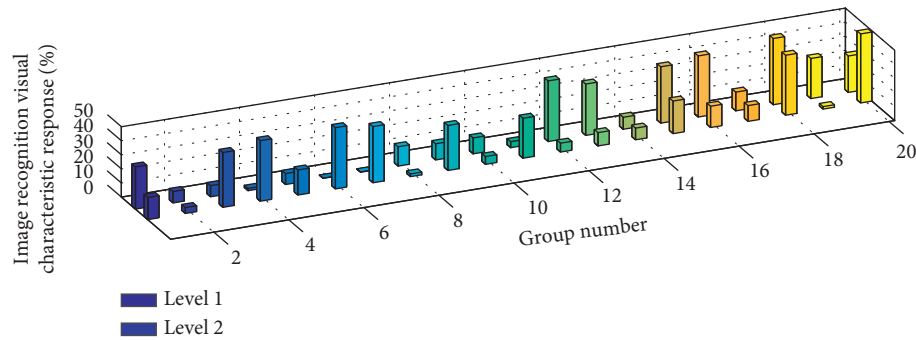
FIGURE 8: Image recognition visual feature frequency response histogram.

second neuron is firing, which means that the identified number is 1. When the threshold is unchanged, we change the film time constant between 5 and 20, and the accuracy fluctuates up and down. When the threshold is between 0.2 and 1.8, the accuracy fluctuates greatly. It can be seen that the threshold is one of the main parameters that affect the accuracy of network recognition. When the threshold value is 1.5, the accuracy reaches the highest value of 93.77%.

5. Conclusion

This paper proposes a dynamic administrative management construction model based on the cloud computing network, which uses the cloud computing network as the intermediary of the connection and the decision-making body for the construction of administrative management. When choosing, the good classification performance of the evolution cloud computing network is applied, and the resource capacity of the Internet of Things can be selected. According to the matching of information and order requirements, the most suitable node Internet of Things is selected to form an optimal administrative management for the corresponding order. The system uses the traditional image recognition algorithm (PCA) and the cloud computing environment to meet the computing efficiency requirements of users for mass image recognition. Certain improvements take the use of image classification algorithms that assign weights and dynamic k values to image images to improve recognition accuracy and stability. Finally, the experiment shows that the system meets the user's requirements for the efficient calculation efficiency and recognition accuracy of the image recognition system. When modeling a dynamic supply chain with time-lag characteristics, applying the orthogonal neural network method can effectively improve the stability of the model, the speed of convergence, and the ability to approximate the optimal solution. The organizational system and structure of dynamic administrative management need to be built on a certain integrated operation platform (including knowledge/skill network, information network, logistics network, and sports network), and the basic platform for dynamic administrative management is also used to a large extent. The above stipulates the specific operation model of dynamic administrative management. Finally, using MatLab to simulate the model, the results show that the dynamic administrative management proposed in this paper has better performance than static administrative management. It can

not only reduce costs but also meet more demand orders. The two-dimensional incentive mechanism can effectively stimulate the overall efficiency of administrative management.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this article.

Acknowledgments

References

[1] L. M. Dang, M. J. Piran, D. Han, K. Min, and H. Moon, "A survey on internet of things and cloud computing for healthcare," *Electronics*, vol. 8, no. 7, p. 768, 2019.

[2] R. K. Barik, H. Dubey, C. Misra et al., "Fog assisted cloud computing in era of big data and internet-of-things: systems, architectures, and applications," *Cloud Computing for Optimization*, pp. 367–394, Springer, Berlin, Germany, 2018.

[3] G. Aceto, V. Persico, and A. Pescapé, "Industry 4.0 and health: internet of things, big data, and cloud computing for healthcare 4.0," *Journal of Industrial Information Integration*, vol. 18, Article ID 100129, 2020.

[4] G. Manogaran, R. Varatharajan, D. Lopez, P. M. Kumar, R. Sundarasekar, and C. Thota, "A new architecture of internet of things and big data ecosystem for secured smart healthcare monitoring and alerting system," *Future Generation Computer Systems*, vol. 82, pp. 375–387, 2018.

[5] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[6] J. Jasperneite, T. Sauter, and M. Wollschlaeger, "Why we need automation models: handling complexity in industry 4.0 and the internet of things," *IEEE Industrial Electronics Magazine*, vol. 14, no. 1, pp. 29–40, 2020.

[7] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of internet of things and cloud computing on healthcare systems: opportunities, challenges, and open problems," *Journal of*

Ambient Intelligence and Humanized Computing, vol. 10, no. 10, pp. 4151–4166, 2019.

[8] Y. Guo, N. Wang, Z. Y. Xu, and K. Wu, "The internet of things-based decision support system for information processing in intelligent manufacturing using data mining technology," *Mechanical Systems and Signal Processing*, vol. 142, Article ID 106630, 2020.

[9] M. Kang, E. Park, B. H. Cho, and K. S Lee, "Recent patient health monitoring platforms incorporating internet of things-enabled smart devices," *International Neurourology Journal*, vol. 22, no. 2, pp. S76–S82, 2018.

[10] A. Malik and H. Om, "Cloud computing and internet of things integration: architecture, applications, issues, and challenges," *Sustainable Cloud and Energy Services*, pp. 21–24, Springer, Berlin, Germany, 2018.

[11] A. Kumari, S. Tanwar, S. Tyagi, N. Kumar, M. Maasberg, and K.-K. R. Choo, "Multimedia big data computing and internet of things applications: a taxonomy and process model," *Journal of Network and Computer Applications*, vol. 124, pp. 169–195, 2018.

[12] A. Sharif, J. P. Li, and M. A. Saleem, "Internet of things enabled vehicular and ad hoc networks for smart city traffic monitoring and controlling: a review," *International Journal of Advanced Networking and Applications*, vol. 10, no. 3, pp. 3833–3842, 2018.

[13] T. Pasquier, J. Singh, J. Powles, D. Eyers, M. Seltzer, and J. Bacon, "Data provenance to audit compliance with privacy policy in the internet of things," *Personal and Ubiquitous Computing*, vol. 22, no. 2, pp. 333–344, 2018.

[14] M. Anuradha, T. Jayasankar, N. B. Prakash et al., "IoT enabled cancer prediction system to enhance the authentication and security using cloud computing," *Microprocessors and Microsystems*, vol. 80, Article ID 103301, 2021.

[15] M. Elhoseny, K. Shankar, S. K. Lakshmanaprabu, A. Maseleno, and N. Arunkumar, "Hybrid optimization with cryptography encryption for medical image security in internet of things," *Neural Computing and Applications*, vol. 32, no. 15, pp. 10979–10993, 2020.

[16] L. Bittencourt, R. Immich, R. Sakellariou et al., "The internet of things, fog and cloud continuum: integration and challenges," *Internet of Things*, vol. 3-4, pp. 134–155, 2018.

[17] G. Manogaran, D. Lopez, C. Thota, K. M. Abbas, S. Pyne, and R. Sundarasekar, "Big data analytics in healthcare Internet of Things," *Innovative Healthcare Systems*, pp. 263–284, Springer, Berlin, Germany, 2017.

[18] M. Maksimović and V. Vujović, "Internet of things based e-health systems: ideas, expectations and concerns," *Distributed Computing in Smart Healthcare*, pp. 241–280, Springer, Berlin, Germany, 2017.

[19] C. Zhang, "Design and application of fog computing and internet of things service platform for smart city," *Future Generation Computer Systems*, vol. 112, pp. 630–640, 2020.

[20] S. Byrne, "Remote medical monitoring and cloud-based internet of things healthcare systems," *American Journal of Medical Research*, vol. 6, no. 2, pp. 19–24, 2019.

[21] I. U. Din, M. Guizani, S. Hassan et al., "The internet of things: a review of enabled technologies and future challenges," *IEEE Access*, vol. 7, pp. 7606–7640, 2018.

[22] X. Gong, L. Duan, X. Chen, and J. Zhang, "When social network effect meets congestion effect in wireless networks: data usage equilibrium and optimal pricing," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 449–462, 2017.

Research Article

Analysis of an Online English Teaching Model Application Based on Improved Multiorganizational Particle Population Optimization Algorithm

Fang Wang (ID)

Jiangsu Lianyungang Higher Vocational and Technical School of Traditional Chinese Medicine, Lianyungang, Jiangsu 222007, China

Correspondence should be addressed to Fang Wang; 1997050027@lygtc.edu.cn

This paper uses an improved multiorganizational particle population optimization algorithm to conduct an in-depth analysis and study of an online English teaching model and uses the altered model for practical applications. The model building elements are extracted from it for the initial construction of a blended learning model of English-speaking teaching in junior high school. The main purpose of the first round of action research is to test the rationality of each element of the model, the main purpose of the second round of action research is to refine the model links and improve the operability of the model, and the main purpose in the third round of action research is to test the perfected model and explore the model. The main purpose of the third round of action research is to test the refined model and explore the application suggestions of the model. After the three rounds of action research, we finally obtained a more mature blended learning model for teaching English as a foreign language in junior high school. Mainly through the comparison of the pre- and posttest scores of English speaking of the experimental subjects and the comparison of the pre- and posttest data of the relevant questionnaires, the following experimental conclusions were drawn: adopting the blended learning-based speaking teaching model can effectively improve students' interest in learning English, their attitudes and their English speaking skills including pronunciation, phonetic intonation, conversational communication, and oral expression and can enhance students' group cooperation and communication ability, independent learning ability, evaluation awareness, and ability. This single-guided learning mechanism can effectively avoid the shocks that are easily caused by the dual-guided role of traditional PSO. The dimensional learning strategy constructs a learning paradigm for each particle by learning from each dimension of the individual optimal position of the particle to the corresponding dimension of the group optimal position, respectively. Dimensional learning is formally integrated into the learning paradigm only if it can improve the fitness of the paradigm so that the dimensional learning strategy can avoid the phenomenon of degradation of the learning paradigm and the phenomenon of "two steps forward, one step back." In the dimensional learning strategy, since each particle learns from best, although it has a strong exploitation capability, it may cause all particles to converge to best quickly, making the algorithm converge prematurely.

1. Introduction

With the development of information technology, traditional classroom teaching can no longer meet the diverse needs of teachers and students in terms of time and location, while online teaching breaks the time and space constraints of traditional classroom teaching and creates an environment where students can learn even outside the classroom.

For this reason, the addition of information technology has led to dramatic changes in the format of teaching, the types of teaching resources, and student learning behaviours. A new teaching model that combines the advantages of online teaching and traditional offline teaching, the hybrid teaching model, has emerged [1]. Offline teaching and online teaching have their advantages and limitations and cannot replace each other. Hybrid teaching mode is a teaching model that

combines the advantages of traditional offline teaching and emerging online teaching [2]. This model makes educational resources richer and learning activities more diverse, especially with the popularity of mobile terminals such as smartphones and iPods and the rapid development of some excellent APPs, which allows students to learn conveniently and quickly anytime and anywhere with a single click on their mobile phones, breaking the limitations of time and space for students in traditional teaching. In recent years, the smart campus is being rapidly built and promoted, education data is growing exponentially, and education is rapidly moving into the era of big data [3]. Teaching is a dynamic, interactive, and continuous process, which is bound to generate a large amount of timely and diverse data. This data serves as a basis for judging student learning, which in turn responds to teachers' teaching standards. Therefore, mining and analysing the data is necessary. Learning analytics is precisely the measurement, collection, analysis, and reporting of educational data, using different techniques to analyse learning data to discover student characteristics and understand teaching and learning. Common techniques include statistical analysis, cluster analysis, and correlation analysis. With the popularity of smart campuses, increased educators and researchers are coming to focus on students' learning behaviours. With the development of information technology in education, learners have increased ways to acquire knowledge and have higher requirements for education, with more focus on whether teachers can tailor and personalize their teaching. Personalized teaching is a new form of teaching that has emerged in recent years, and it is a way to achieve good teaching results by stimulating learners to think and develop [4]. The teaching process will be optimized for the needs of each learner and is closely related to the learner's interests. In the context of education informatization, education big data, and personalized teaching, the blended teaching model has emerged. It is of strong practical significance to analyse students' learning behaviour under the blended teaching model. Dimensional learning can only be formally integrated into the learning paradigm on the premise that it can improve the fitness of the paradigm. Therefore, the dimensional learning strategy can avoid the phenomenon of learning paradigm degradation. To gain a deeper understanding of the teaching of blended teaching mode and analyse students' learning behaviours under the blended teaching model, this study classifies students with the characteristics of learning behaviours, explores the correlation between learning behaviours and academic performance, and identifies the factors affecting learning behaviours. It is applied in practice to make up for the lack of research on learning behaviour in blended learning and teaching.

Particle swarm optimization algorithms are the primary choice for optimization problems due to less memory required, easy implementation, faster convergence, and the ability to provide better performance on different benchmark test functions and engineering problems. Particle swarm optimization algorithms are population-based iterative algorithms and the requirement of many evaluation functions for computation makes particle swarm

optimization algorithms inappropriate for optimization areas where fitness computation is time-consuming. In solving complex problems with multiple extremes, the standard particle swarm algorithm tends to fall into local extremes, resulting in premature convergence without finding the optimal value of the entire search space; when the dimensionality of the search space is high, the algorithm has poor search capability [5]. These drawbacks have limited the wider application of particle swarm optimization algorithms. Therefore, speeding up convergence and avoiding getting into local extremes have become two of the most important and attractive goals in the research of particle swarm optimization algorithms. Exploration and exploitation are the two main search mechanisms in the search for optimal solutions in particle swarm optimization algorithms. To achieve fast convergence and avoid getting into local extremes, one must balance the exploration and exploitation capabilities of the population [6]. Exploration searches the region of the search space with potentially optimal solutions; more refined searches performed in the vicinity of the population optimal solution. Overemphasis on exploration will lead to a decrease in convergence by wasting search time in nonoptimal regions. On the other hand, overemphasis on exploitation will reduce the diversity of the population in the search process, which may be the population caught in the local optimal region. Therefore, how to balance the exploration and exploitation capabilities in the search process is a question worthy of deeper investigation.

In this paper, a two-population particle swarm optimization algorithm with a dimensional learning strategy is proposed from the perspective of effectively discovering and preserving good information in each dimension of the population, and the ability of the proposed algorithm to balance exploration and exploitation intensity is discussed. The particle swarm optimization algorithm is a heuristic stochastic search algorithm, and this stochastic search mechanism lacks certainty and does not guarantee that the algorithm can necessarily find the optimal solution. Online learning platforms provide students with rich resources, such as text, images, audio, video, and animation, but these resources mostly presented directly, the organization of resources is confusing, and the content and quality need to be further optimized. This leads to the phenomenon of students' knowledge being lost in the face of numerous learning resources. Again, there is a lack of relevance and adaptability. Although student autonomy has been increased, the content and path of learning remain neat and uniform for everyone, and the test questions are fixed.

1.1. Status of Research. How to break the shackles of traditional teaching methods and achieve the personalized development of students has become a major dilemma facing education in China at present [7]. With the development of big data, cloud computing, and artificial intelligence technology, the way of learning, teaching, and cognition is undergoing radical changes. Especially the development of big data in education makes the collection

and deep analysis of educational data possible and makes the sharing of educational resources a reality [8]. How to be able to collect and integrate data and use the data to discover the hidden behind it becomes an important way to enhance the core competitiveness of education, and the whole society has entered the era of big data, helping particles escape from local extrema. TSLPSO reduces the pressure of DLPSO adaptability assessment by introducing dimensional learning subgroups. The education field has set off a boom of promoting education reform and development based on big data in education, and the research on big data has shown an unprecedented development momentum [9]. At the level of applied research, a personalized adaptive learning platform based on big data has become the focus of research. Adaptive learning platforms collect student interaction data based on big data technology and dig deeper and analyse them for learning, providing students with personalized learning services according to the individual differences of learners, which have taken root in many countries and developed into personalized adaptive learning platforms [10]. The US Adaptive Learning Platform is one of the most mature adaptive learning platforms at present. After decades of research, it integrates the research results of psychology, measurement, cognitive learning theory, and intelligent learning systems and pioneers the design and application of personalized services for big data in education. By collecting students' online learning data, it can accurately predict and analyse students' strengths, weaknesses, learning interests, cognitive levels, and participation levels and personalize learning contents. It not only meets the needs of students to a great extent but also brings great convenience to teachers, parents, and school administrators in assessment and management [11]. At this stage, China has also made many useful attempts in the theory and practice of adaptive learning, but most of the research remains at the theoretical level, and there are few applications and practices in this area, and there is no mature adaptive learning platform yet. Therefore, through the study of Newton's adaptive learning platform in the United States, it can inspire the development of adaptive learning platform and personalized teaching in China.

The neural network uses a multilayer structure where the basic processing unit is a neuron in each layer of the network and contains both linear superposition and nonlinear activation as two types of computation, so mathematically speaking, a neural network is a structure with nesting that can be used to describe highly and nonlinear mapping relationships [12]. Neural networks are not all smooth sailing, it has experienced many troughs since it was proposed, but each time it came out of the trough was when it made a breakthrough and gradually became the core algorithm of current artificial intelligence [13]. Its main feature is the backpropagation algorithm, thus significantly reducing the time required for model training. Subsequently, scholars have never stopped their research on BP neural networks. Because it did not have the problem of gradient disappearance like the neural network that used more layers of the network in the environment with a relatively small amount of data at that time, this

algorithm with the characteristics of simplicity and efficiency, fast training speed, and applicable to small sample size once overtook the deep neural network to become the mainstream of machine learning.

It has attracted many scholars and researchers to conduct in-depth research on it, thanks to its relatively simple concept, relatively easy implementation, and better global search capability when solving some more complex optimization problems, such as when the objective function is multipeaked, nonlinear, and nonderivative. Nowadays, the PSO algorithm has been applied to many engineering practices such as pattern recognition, intelligent robotics, and signal processing. The improved algorithm is used to solve the selection of parameters in the training process of machine learning, to improve the model performance and make the BP neural network in the initialization stage be in a relatively low loss for training, and to achieve the training requirements faster and better model and also can make the SVM model have better hyperparameters and the model after training have better results.

2. Improved Multiorganizational Particle Population Optimization Algorithm for Online English Teaching Model Applications

2.1. Improved Multiorganizational Particle Population Optimization Algorithm Design. The topology of a population defines the way information is shared and information has interacted between particles. Particle swarm algorithms based on neighbourhood topology control the exploration and exploitation capabilities of the algorithm according to the different information-sharing mechanisms between particles. A new information flow mechanism is proposed to update the position of each particle, namely, the full information particle swarm optimization algorithm (FIPS) [14]. The FIPS algorithm uses a weighted average of the individual optimal positions of all neighbouring particles to update the position of a given particle. The unified particle swarm optimization (UPSO) algorithm is proposed, which uses the optimal experience from both local and global neighbours to control the exploration and exploitation capabilities of the population. A dynamic neighbourhood learning particle swarm optimization (DNLPSO) algorithm is proposed, where exemplar particles are selected from the optimal positions of neighbouring particles including itself so that the velocity of a particle may be influenced by both the historical experience of its neighbours and its own historical experience; when the population perceives that the search has stalled, the population changes the flow of information. When the fitness of particle i does not improve within K generations, the particle is judged to be in stagnation, at which point the particle updates its neighbours by linking a new particle [15]. The particle uses a betting round selection strategy (based on particle fitness ranking) to select a new neighbour. In this way, underachievers who lack practice must complete as many as eight question sets or more, while students with stronger learning ability may complete adaptive follow-up after one question set.

Analysis of the PSO algorithm formulation shows that the particle flight trajectory is influenced by its own flight experience (individual optimum) as well as the global optimum information in the early stages and converges to the global optimum in the later stages [16]. This suggests that there is a better global optimum information point that will be able to guide the particle towards a better solution space. To obtain better global optimum information, the HPSO algorithm is proposed to achieve the position update by performing multiple Corsa mutations on the optimal particles. Most of the current mutation strategies for optimal particles pick all dimensions or randomly pick some dimensions for mutation; for high-dimensional complex functions, the computation of their fitness will be inefficient due to the interference between dimensions, which causes some dimensions to become better but be masked by other dimensions that become worse; compared to multiple mutations, the dimension-by-dimension mutation is more efficient. The obtained variational solutions are usually better [17]. Based on this, a variational strategy of dimension-by-dimension backward learning of the centre of gravity for the best particles is proposed, where dimension-by-dimension variation can reduce interdimensional interference and backward learning of the centre of gravity can expand the search space, increase the diversity of populations, and improve the convergence accuracy. However, the computational overhead of dimension-by-dimension backward learning is large, and it is not suitable for all particles to be mutated dimension by dimension:

$$Gbest_i = 3 \times k \times M_i + Gbest_j. \tag{1}$$

With the advent of the network era and the transformation of learning tools, methods, and resources, a more suitable connectivism for the digital information age is proposed based on behaviourism. Among them, connectionist learning theory believes that the starting point of all knowledge and learning is the individual, and the individual's knowledge can form a network, and in the process of learning and sharing knowledge, the nodes of the connection network and the nodes of other individuals, organizations, or institutions collide and interact with each other, and the connection of old and new nodes makes the continuous expansion of the individual's original network form a larger knowledge network. In the process of online collaborative learning, learners no longer exist as individual individuals but form learning groups through their will and rational choice and complete teaching tasks with learning groups as the basic unit. The learner's knowledge network collides and interacts with the knowledge networks of other members of the learning group, jointly expanding the original knowledge network, improving the learner's knowledge framework system, and enhancing the effect of online collaborative learning, as shown in Figure 1.

From the above example, we can see that when the dimensional learning strategy constructs a learning paradigm, each time the dimensional status of the temporary paradigm is updated, it is compared with the current learning paradigm, and if it is better than the current learning paradigm, the learning paradigm is updated to the current temporary paradigm; otherwise, the current learning paradigm is not changed and the learning process continues for the next dimension. The effect of face-to-face oral English teaching has been greatly improved. After class, dialogue exercises are arranged through the wing class network, and dialogue exercises are continued through the WeChat group cooperation group to achieve language review, consolidation, extension, and expansion and train students to speak and use English after class, which are good speaking habits:

$$Div = \frac{1}{n}\sum_{i=1}^{m}\sqrt{\sum_{j=1}^{m}\left(x_{i,j}+x_j\right)^2},$$

$$x_j = \frac{\sum_{j=1}^{m}\left(x_{i,j}+x_j\right)^2}{n}. \tag{2}$$

We give plots of single-peaked and multipeaked functional diversity. The dimensionality of the search space in the diversity validation experiment is 30 dimensions, the function is evaluated 300,000 times, and the population size is 20. It is seen that the integrated learning subpopulation diversity is higher than the diversity of the dimensional learning subpopulation, and the diversity of the whole population lies between the two subpopulation diversities. The results from the diversity show that the dimensional learning subpopulation maintains a smaller population diversity and therefore converges quickly. As expected, the integrated learning subpopulation maintains a higher group diversity because there is no information as a central guiding direction, so the group does not converge to a smaller extent quickly. The balance of exploration and exploitation provided by the dimensional learning subgroup in collaboration with the integrative learning subgroup resulted in an intermediate diversity ranking for the whole group, indicating that the integrative learning strategy we introduced did increase the diversity of the group. Thus, the results of the diversity comparison validate the design expectation that the dimensional learning subgroup tends to local exploitation, while the integrated learning subgroup tends to global exploration. The interaction and cooperation of the two subpopulations ensure global search and fast convergence of the population:

$$f_7(x) = \sum_{i=1}^{n}\left[x_i^2 + 10\,\sin\left(2\tau x_i\right) - 10\right],$$

$$f_{16}(x) = \sum_{i=1}^{n}\left[a_i x_i^2 - 10\,\sin\left(2a_i\tau x_i\right) + 10\right]. \tag{3}$$

It is difficult to find its optimal solution for most of the algorithms, and when the dimension of the search space is higher than 3, it can be considered as a multipeaked function (Figure 2). Therefore, most of the algorithms have unsatisfactory results on functions. Although the optimal solution for the 3 functions out of 5 single-peaked functions is found, one more than the proposed one, it performs the worst on the function. Even though only the dimensional learning strategy was used, it still achieved the optimal result on the noise function. Overall, the algorithm shows better robustness on the single-peaked functions (i.e., the first 5 benchmark functions)
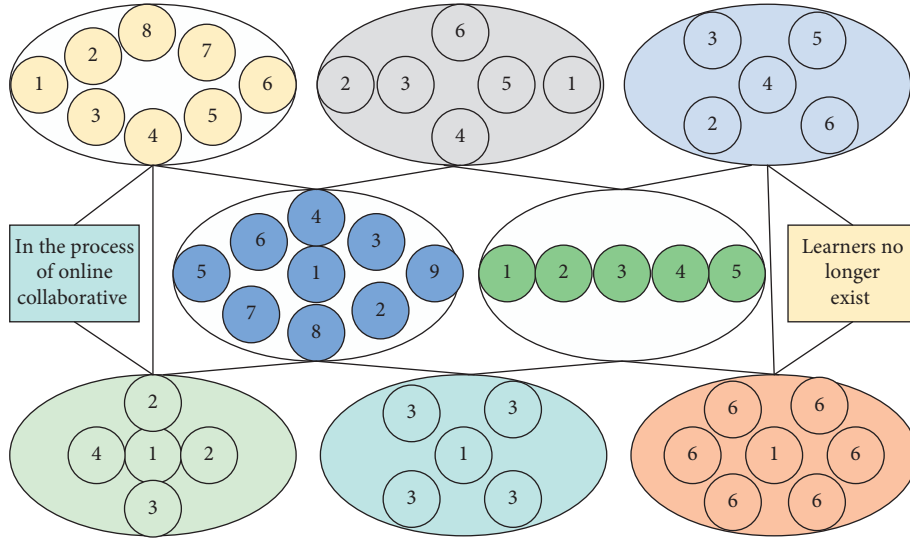
FIGURE 1: Framework of improved multiorganizational particle population optimization algorithm.
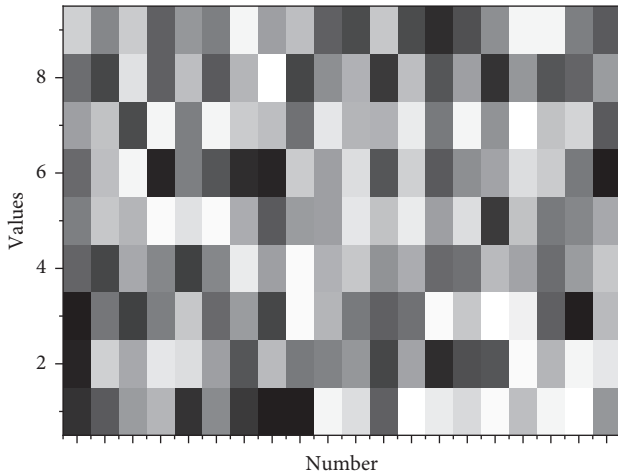
FIGURE 2: Contour plot of the function.

and thus obtains a smaller sum than the champion algorithm. Multipeaked functions contain multiple locally optimal solutions, which may lead to premature convergence of the PSO algorithm. It is difficult to find the global optimal solution of the function with conventional PSO because the problem has many deep local optima far from the global optimal solution. The nodes of organizations or institutions collide and interact with each other, and the connection of new and old nodes makes the original network of individuals continue to expand and form a larger knowledge network. Once a particle of classical PSO is trapped in a deep local optimum, it is difficult to escape. The improved multiorganizational particle population optimization algorithm has better resistance to the local optimum trap and thus achieves a high accuracy of 3.82E-04. The other 5 bodies, however, cannot converge to the global optimum effectively, and the worst result is 2.38 E+03. Functions are very complex multipeaked functions with many local extrema. For this problem, algorithms that maintain better diversity tend to produce better results. The algorithm achieves

the optimal solution on both functions, 0. This is due to the dimensional learning strategy and the integrated learning strategy used in the algorithm TSLPSO, which effectively enhances the diversity of the particle search while improving the convergence accuracy.

The users have different rating values for the target object, so a weighted average of the ratings is taken as the prediction result, and the weights can be obtained from the similarity between similar users and the target user:

$$R_{ik} = \mu_{ui} - \frac{\sum_{i,j} P(U_i U_j)}{\sum_{i,j} P(U_i^2 U_j^2)^2}. \tag{4}$$

Matrix decomposition is the decomposition of a matrix into the product of two or more matrices. Matrix decomposition-based recommendation algorithms are used to build user rating prediction models by modelling known user ratings of objects, which are mainly applied to two scenarios: rating prediction and TOP-N recommendation. Currently, the widely used matrix decomposition-based recommendation algorithm is the number of hidden factor models (LFM). The basic idea of LFM is to link user interests and items through implicit features, that is, the existence of similarity. The earliest appearing hidden factor model is the singular value matrix decomposition model, referred to as SVD, which is calculated as

$$R_{M \times N} = U_{M \times N} \sum M(M \times N). \tag{5}$$

According to Wittrock, the king of generative learning theory, the essence of learning is a process of active construction and generation of meaning, in which learners construct meaning by actively establishing relationships between external stimuli and their original cognition. This construction of "special associations between external information and existing knowledge and experience" is called generation. The basic assumption of generative learning theory is that "the human brain is not a passive recipient of

information, but an active constructor" [18]. In other words, the learner is not a passive recipient, but an active participant in the learning process, actively constructing a meaningful understanding of the information in the surrounding environment. However, without attention, motivation, and existing knowledge in memory, the construction of relationships between existing knowledge and new information will not occur. The comprehensive learning subgroup maintains a high group diversity, because there is no information from the central guide, so the group will not quickly gather into a smaller area. By assimilating Plus's information processing theory, Wittrock proposes the information processing process of the generative learning model: attention, motivation, prior knowledge experience, and generation. Attention is the directional factor that guides the generative process, which points the generative process to prior knowledge and experience. Motivation refers to the desire to actively generate these two connections and to attribute the effectiveness of generating them to the degree of one's effort; prior knowledge and experience include existing concepts, reflective cognition, abstract knowledge, and concrete experience, and the learner's prior cognitive structure and level are extremely important in generating the meaning of things; generation refers to the intrinsic connections that form new knowledge and the connections between new knowledge and existing experience.

2.2. Design of Online English Teaching Model Applications.
Constructivism assumes that students' knowledge is constructed based on their prior knowledge experiences and cognitive structures and is actively acquired through meaningful construction. The meaningful construction of knowledge requires interaction with the external environment and the use of other aids from certain situations, for example, with the help of other people including teachers, classmates, and peers, using the necessary learning materials and learning resources. Constructivist learning theory emphasizes the cognitive dynamics, the contextual nature of learning, the importance of resources for the construction of meaning, and the design of the teaching environment. The learning-based instructional design theory is developed in response to the above requirements of constructivist learning. It provides new ideas and new ways of designing and transmitting teaching and learning to overcome the one-way transmission, passivity, and closure in traditional instructional design. This instructional design theory is particularly clear in the diagrammatic representation of the constructivist learning environment model proposed by David Jonathan, a contemporary advocate, and promoter of constructivist instructional design, as shown in Figure 3.

Blended learning has teacher-led student-led as its core idea, and both teacher teaching and student-led learning are extremely important. Constructivist learning theory also emphasizes the dual role of teaching and learning, which has a guiding role in the design of blended learning models. At the same time, constructivism attaches importance to the design of the teaching environment, and blended learning has rich multimedia and network technologies that can
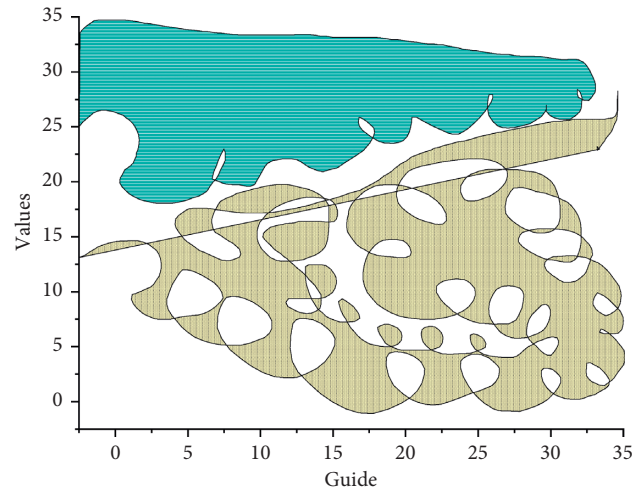
Figure 3: Diagrammatic representation of the constructive learning environment design model.

provide material support for the learning environment required by constructivism. Through the guidance of constructivist theory, the teaching objectives of blended learning can be effectively achieved and students' abilities in all aspects can be developed [19]. Once the particles of the classic PSO fall into the deep local optimum, it is difficult to escape. It has better ability to resist local optimal traps, so it has achieved a high precision of 3.82E-04. The researcher believes that with the development of mobile Internet technology and the improvement of information delivery tools, constructivist learning theory will continue to be further developed as the concept of blended learning is proposed and practiced and provides a broader space for putting it into practical application.

When teaching English as a foreign language in a blended learning environment, the first step for teachers is to present speaking tasks. Driven by the tasks, students can use their minds to use English actively by completing specific tasks and actively participate in various tasks to truly "learn by doing" and gain and accumulate corresponding learning experiences and enjoy the fun of learning. In the practical application of oral English teaching, task-based teaching makes students immerse themselves in the situation, truly gain emotional experience and adjust learning strategies, form a positive learning attitude, and promote the improvement of their practical language skills. The flexible use of cooperative learning in blended English speaking teaching activities can create a classroom atmosphere, stimulate students' interest in learning English, overcome classroom anxiety, and motivate students to actively exert their subjective initiative or maintain a high level of participation in online learning, increase the opportunities for students to use English to communicate and complete speaking learning tasks together, and improve students' speaking learning performance and speaking communicative competence.

Multiobjective optimization problems exist with multiple optimization objectives, and these subobjectives are often independent and conflicting with each other and cannot reach the optimal solution simultaneously.

Theoretically, most optimal grouping processes are multiobjective problems, and solutions should be designed in a multiobjective framework. Based on the educational theory of collaborative learning, the basic principle of learning group formation is "intergroup homogeneity and intragroup heterogeneity," where intergroup homogeneity ensures the success of all learning groups, while intragroup heterogeneity depends on different research topics and teachers' teaching needs. According to the specific requirements and principles in the proposed model of learning group formation, the multidimensional objective functions are set correspondingly. Due to the characteristics of the problem with many constraints, adding up the constraint violation degrees of all constraints may make the feasible domain of the algorithm smaller, which directly leads to few feasible solutions found and low reliability of the algorithm. Therefore, this study adopts the scheme of treating the constraint violation degree of each constraint as an objective separately, as shown in Figure 4.

In the real world, there do not exist any single-objective problems; in other words, single-objective problems are defined mainly for simplicity. This means that most of the time a person just chooses to consider the most important objective problem and ignores the others, thus converting a multiobjective problem into a single-objective problem. In addition, sometimes just one objective is chosen, and one or more other objectives are considered as constraints [20]. In both cases, the optimization process is simplified and redefined as single-objective optimization. Theoretically, most optimal grouping processes are multiobjective problems, and solutions should be designed in a multiobjective framework. Learning group formation as a multiobjective optimization problem is an essential and complex step in effective collaborative learning, and the purpose of this study is to propose a method based on a heuristic search strategy to enhance the intergroup homogeneity and intragroup heterogeneity of learning groups in collaborative learning environments, capable of grouping any number of prequalified learners with multiple characteristics into an arbitrary number of optimal inter- and intrahomogeneous groups [21].

Therefore, learning groups with smaller total mean deviation percentages are more favourable solutions. Through the information processing theory of absorption and addition, Whitlock proposes the information processing flow of the generative learning model: attention, motivation, prior knowledge and experience, and generation. The previous statement clearly shows that heterogeneity between groups and within groups is mainly reflected by values that reflect the average ability of all groups in a solution. Minimizing the overall error value of the learning group formula does not guarantee that the error values of all groups that make up the learners are minimized simultaneously. In other words, when the optimization algorithm is arranged to evaluate each learning group by its average fitness value, it is set to be insensitive to the fitness value of the constructing learners.

3. Results and Analysis

3.1. Performance Results of Improved Multiorganizational Particle Population Optimization Algorithm. To compare the convergence speed, reliability, and successful performance of the algorithms, Figure 5 gives some key metrics, including the average number of objective function evaluations (FEs), success rate (SR), and success performance (SP) of each algorithm when it successfully achieves acceptable optimization accuracy. FEs and SRs are used to quantify the convergence speed and reliability of the algorithm. From Figure 5, TSLPSO converges fastest on 7 benchmark functions, followed by DLPSO (fastest convergence on 4 functions, i.e., functions F2, F6, F7, and F14), GL-PSO (fastest convergence on 3 functions, i.e., functions F1, F3, and F10), and L-SHADE (fastest convergence on 2 functions, i.e., functions F4 and F5). From the results of SR, TSLPSO is the most reliable with an average success rate of 100%, followed by DLPSO with a success rate of 98.62%.

From the convergence graph of single-peaked functions, the algorithm TSLPSO shows high convergence accuracy and fast convergence speed on most of the single-peaked functions except for the function F5. The convergence graphs of the multipeaked functions show that the algorithm TSLPSO not only converges with high accuracy but also converges faster than most of the algorithms. Moreover, DLPSO with only a dimensional learning strategy shows very competitive performance on these multipeaked functions. Thus, TSLPSO is generally effective for both single-peaked and multipeaked functions in terms of accuracy and convergence speed.

Most existing PSO algorithms use a linear weighting of individual optimal positions, random combination by dimension, or orthogonal combination of individual optimal positions and group optimal positions to construct learning examples, which guide the particle search instead of individual optimal solutions and group optimal solutions in the classical PSO velocity update formulation. This single-guidance learning mechanism can effectively avoid the oscillation phenomenon that is easily generated by the traditional PSO double-guidance action.

However, the learning paradigms constructed by the above methods are highly stochastic and cannot ensure that the learning paradigm will not degenerate even when the individual optimal solution and the population optimal solution evolve every generation. Particles learning from degenerate paradigms is not conducive to maintaining the efficiency of the algorithm. The dimensional learning strategy proposed in this paper allows each dimension of the individual optimal position of the particle to learn from the corresponding dimension of the population optimal position respectively when constructing the paradigm, and the dimensional learning is formally incorporated into the learning paradigm only if it can improve the fitness of the paradigm, so that the dimensional learning strategy can avoid the phenomenon of degenerating the learning paradigm and the phenomenon of "two steps forward, one step back" phenomenon, as shown in Figure 6.
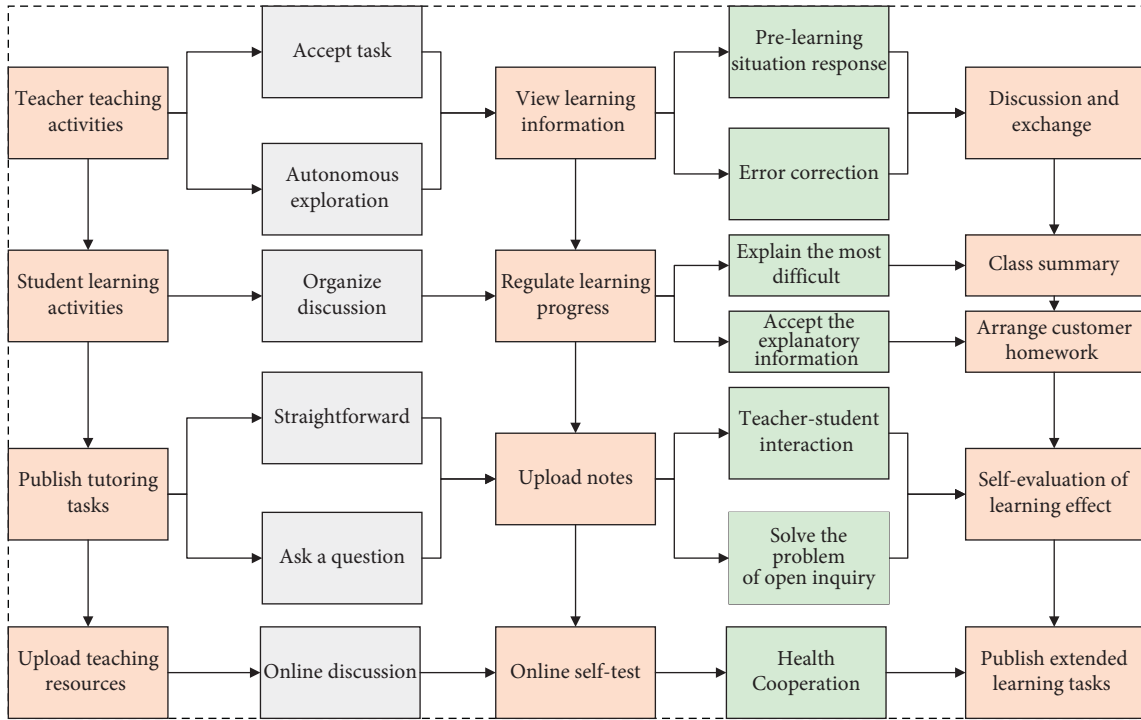
Figure 4: Model application design model.
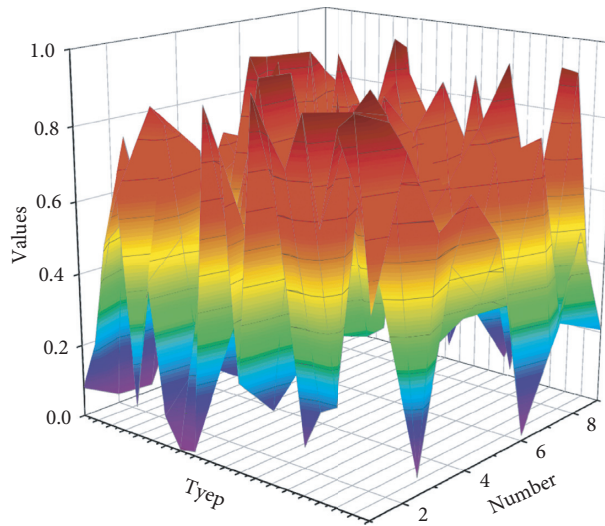
Figure 5: Comparison of the results of the tested algorithms on the benchmark function.

Although the learning paradigm constructed by dimensional learning can guide particles to search towards better regions, the fact that most of the particles are close to the population optimal position may cause premature convergence problems. To address this problem, we introduce an integrated learning strategy to enhance population diversity and help particles escape local extremes. TSLPSO alleviates the pressure of DLPSO adaptation evaluation by introducing dimensional learning subpopulations. Learners exhibit differences in learning behavior characteristics across resource presentations when learning online. In exploring document-based learning resources, this subsection focuses on three dimensions of learners' learning characteristics, namely, different learning knowledge types, course resource hotness, and course resource learning attainment, and the data visualization results are shown in Figure 6. During the study, the courseware resources that do not meet the procedural and descriptive knowledge characteristics are manually screened out, and only the courseware resources of these two major knowledge types are retained as the study objects.
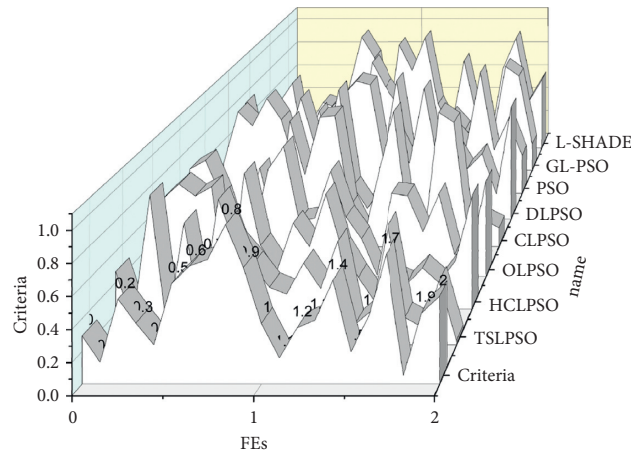
FIGURE 6: Graph of problem convergence curve.

3.2. Results of Applying the English Online Teaching Model.

Adaptive follow-up assignments are personalized assignments on the Newton platform that take advantage of the instructional moment after proficiency learning to correct students' errors before they develop a basic understanding and before they move on to the next topic of study. Student work is completed to determine what students have mastered and understood and what they have not, and questions are selected from the Mastering Item Library to close individual gaps in understanding. Adaptive follow-up assignments also analyse student performance on proficiency learning tasks to select the types of assignments that are most helpful to students. Adaptive follow-up assignments are staged, with each set of questions based on the results of the previous set of questions. Promote the improvement of their language practical ability, and the flexible use of cooperative learning in mixed oral English teaching activities can create a classroom teaching atmosphere, stimulate students' interest in learning English, and overcome classroom anxiety. Students perform adaptive follow-up assignments until they complete the assigned problem set or master all concepts and material. In this way, students who lack practice and are late learners are required to complete as many as eight or more problem sets, while more advanced learners may be able to complete adaptive follow-up after one problem set. After the student completes the adaptive follow-up assignment, the student can see which questions he or she answered and how well he or she answered them. Adaptive follow-up assignments can be added to all proficiency learning tasks.

The teacher first sets the number of problem sets the student receives, each of which takes about 15 minutes, but the specifics will vary from person to person; after this, a total point value is assigned to the assignment, and finally, a deadline is set and whether the student who performs well is expected to automatically receive full credit without taking the adaptive follow-up assignment. Adaptive follow-up assignments provide students with a personalized learning experience that helps each student continue to make progress in the course at the most appropriate time and in the best way possible. Overall, Newton's assignment resources can be categorized as fully adaptive assignments and partially adaptive assignments, as shown in Figure 7.

From Figure 7, it can be seen that in the resource hotness dimension, the mean of video resources above 10 min is higher than the mean of video resources between 5 and 10 min, and the mean of video resources between 5 and 10 min is higher than the mean of video resources between 0 and 5 min. Video resources above 10 min have a higher hotness H but a lower mean of resource learning attainment G, indicating that learners did not watch the whole video. The video resources of 0–5 min have a higher mean value than the mean value of course resource learning attainment G, although the hotness is lower, indicating that learners repeatedly watch the video resources. The resource designer can adjust the length of the video according to the purpose; for example, if the purpose is only to give learners a general understanding of the knowledge, then a video length of more than 10 min or 5–10 min can be used. If the purpose is to give learners knowledge, then a video length of 0–5 min can be used. Depending on the content of the video with or without the teacher and with or without subtitles, the learning characteristics of learners were explored in three dimensions: video resource presentation, course resource hotness, and course resource learning attainment, as shown in Figure 8.

The resources used in this round of action research were mainly topic-related lesson materials as well as microlearning resources. Before face-to-face teaching, teachers posted the resources in a WeChat class group for prelesson viewing. By posting interesting and targeted resources, students' interest in oral learning can be stimulated, and students can focus on the language points before class, so that they can be fully engaged in classroom activities with "excitement" and questions, thus improving students' communicative competence more effectively. Through WeChat's video forwarding function, video resources related to the background of the topic were released for students to watch before class to understand the linguistic situation of the language, which enriched students' language materials and made them understand and think about the topic before class, greatly improving the effect of face-to-face oral
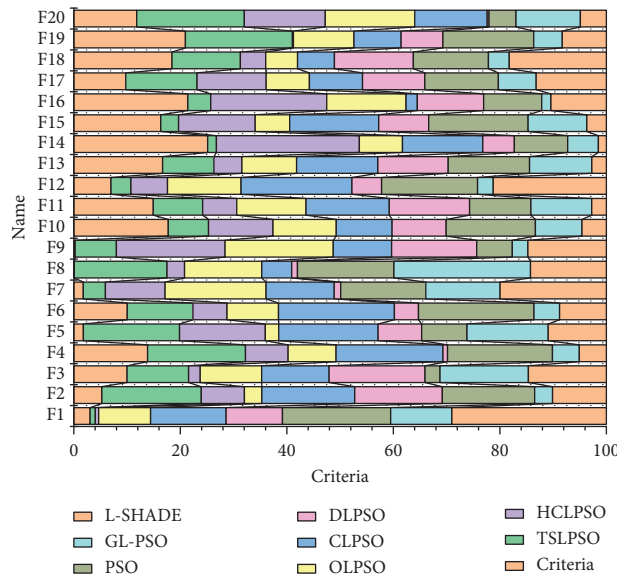
Figure 7: Analysis of video length differences.
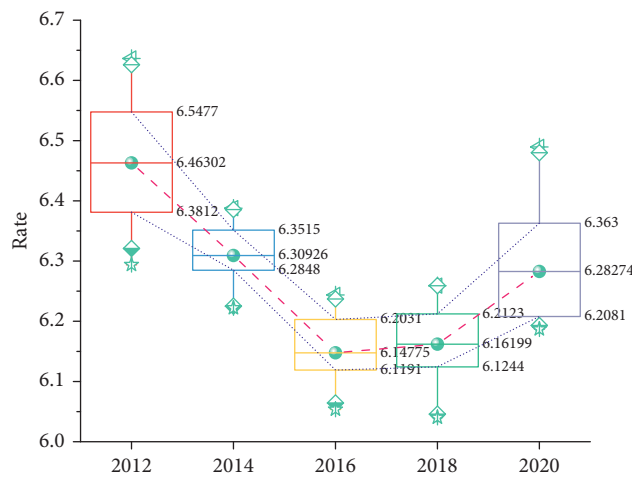
Figure 8: Pre- and postexperimental class results.

teaching. After class, dialogue practice assigned through the wing class network and dialogue drills are continued through the WeChat group cooperation group to achieve language review, consolidation, extension, and expansion and to cultivate students' good speaking habits of speaking and using English after class.

4. Conclusion

In the three rounds of action research, the researcher designed the content and teaching process of oral language teaching for the model and continuously reflected and optimized the implementation strategy and improved the teaching process in the teaching practice. The researcher conducted several questionnaires on the practical application effects of the model in speaking teaching, and after the longitudinal and cross-sectional pre- and posttest data comparison and analysis, it was found that after the

experiment, the students in the experimental class had a greater improvement in their interest and attitude in English learning, cooperative communication ability, independent learning ability, evaluation awareness, and ability, problem-solving effect, and resource application effect in teaching compared to the students in the control class. This leads to the conclusion that using a blended learning-based speaking teaching model can effectively improve students' interest and attitude in English learning, can improve students' group cooperative communication and independent learning ability, can improve students' awareness and ability of evaluation, and can improve the effectiveness of problem-solving and the application of resources in teaching. Combining the DCOPSO algorithm with SVM, the traditional SVM hyperparameter search has the characteristics of poor globalization and low accuracy, while using the DCOPSO algorithm for search can solve these problems. The strategy enhances the diversity of the population and helps

the population to jump out of local extremes. Finally, a two-population learning particle swarm optimization algorithm with heterogeneous learning strategies is proposed based on the dimensional learning strategy and the integrated learning strategy. One subpopulation uses a learning paradigm constructed by the dimensional learning strategy to guide the local search of particles, and the other subpopulation uses a learning paradigm constructed by the integrated learning strategy to guide the global search of particles. The two subpopulations achieve mutual collaboration in the search process through different information interaction mechanisms, which effectively improves the algorithm performance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

There are no conflicts of interest in this paper.

Acknowledgments

References

[1] Z. Sun, M. Anbarasan, and D. Praveen Kumar, "Design of online intelligent English teaching platform based on artificial intelligence techniques," *Computational Intelligence*, vol. 37, no. 3, pp. 1166–1180, 2021.

[2] A. Yadav and D. K. Vishwakarma, "A comparative study on bio-inspired algorithms for sentiment analysis," *Cluster Computing*, vol. 23, no. 4, pp. 2969–2989, 2020.

[3] S. AbdulRahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi, and M. Guizani, "A survey on federated learning: the journey from centralized to distributed on-site learning and beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5476–5497, 2021.

[4] S. Jin, S. Kim, H. Kim, and S. Hong, "Recent advances in deep learning-based side-channel analysis," *ETRI Journal*, vol. 42, no. 2, pp. 292–304, 2020.

[5] X. Kong, S. Tong, and H. Gao, "Mobile edge cooperation optimization for wearable internet of things: a network representation-based framework," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 5050–5058, 2020.

[6] R. Hou, Y. Kong, B. Cai, and H. Liu, "Unstructured big data analysis algorithm and simulation of Internet of Things based on machine learning," *Neural Computing & Applications*, vol. 32, no. 10, pp. 5399–5407, 2020.

[7] Q. Ahmed, S. A. Raza, and D. M. Al-Anazi, "Reliability-based fault analysis models with industrial applications: a systematic literature review," *Quality and Reliability Engineering International*, vol. 37, no. 4, pp. 1307–1333, 2021.

[8] H. Song, C. E. xiu-ying Han, C. E. Montenegro-Marin, and S. krishnamoorthy, "Secure prediction and assessment of sports injuries using deep learning based convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3399–3410, 2021.

[9] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, 2021.

[10] E. Adi, A. Anwar, Z. Baig, and S. Zeadally, "Machine learning and data analytics for the IoT," *Neural Computing & Applications*, vol. 32, no. 20, pp. 16205–16233, 2020.

[11] A. A. Marouf, M. K. Hasan, and H. Mahmud, "Comparative analysis of feature selection algorithms for computational personality prediction from social media," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 3, pp. 587–599, 2020.

[12] J. Sun, A. Tárnok, and X. Su, "Deep learning-based single-cell optical image studies," *Cytometry, Part A*, vol. 97, no. 3, pp. 226–240, 2020.

[13] C. Liu, Y. Feng, D. Lin, L. Wu, and M. Guo, "Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques," *International Journal of Production Research*, vol. 58, no. 17, pp. 5113–5131, 2020.

[14] K. Pradhan and P. Chawla, "Medical Internet of things using machine learning algorithms for lung cancer detection," *Journal of Management Analytics*, vol. 7, no. 4, pp. 591–623, 2020.

[15] A. Onan, "Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 572–589, 2021.

[16] L. Ma and B. Sun, "Machine learning and AI in marketing-connecting computing power to human insights," *International Journal of Research in Marketing*, vol. 37, no. 3, pp. 481–504, 2020.

[17] I. Lee and Y. J. Shin, "Machine learning for enterprises: applications, algorithm selection, and challenges," *Business Horizons*, vol. 63, no. 2, pp. 157–170, 2020.

[18] X. N. Bui, H. Nguyen, and Y. Choi, "Prediction of slope failure in open-pit mines using a novel hybrid artificial intelligence model based on decision tree and evolution algorithm," *Scientific Reports*, vol. 10, no. 1, pp. 1–17, 2020.

[19] A. Sharafati, S. B. Haji Seyed Asadollah, D. Motta, and Z. M. Yaseen, "Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis," *Hydrological Sciences Journal*, vol. 65, no. 12, pp. 2022–2042, 2020.

[20] G. S. Budhi, R. Chiong, I. Pranata, and Z. Hu, "Using machine learning to predict the sentiment of online reviews: a new framework for comparative analysis," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2543–2566, 2021.

[21] A. M. Aleesa, B. B. Zaidan, A. A. Zaidan, and N. M. Sahar, "Review of intrusion detection systems based on deep learning techniques: coherent taxonomy, challenges, motivations, recommendations, substantial analysis and future directions," *Neural Computing & Applications*, vol. 32, no. 14, pp. 9827–9858, 2020.

Research Article

An Internet of Things-Oriented Adaptive Mutation PSO-BPNN Algorithm to Assist the Construction of Entrepreneurship Evaluation Models for College Students

Huaxiang Fu [ORCID]

Zhejiang Industry Polytechnic College, Shaoxing, Zhejiang 312000, China

Correspondence should be addressed to Huaxiang Fu; 20060030@zjipc.edu.cn

In this paper, the IoT-based adaptive mutation PSO-BPNN algorithm is used to conduct in-depth research and analysis of the entrepreneurship evaluation model for college students and practical applications. This paper details the principle, implementation, and characteristics of each BP algorithm and PSO algorithm. When classifying college students' entrepreneurship evaluation based on BP neural network, because BP algorithm is a local optimization-seeking algorithm, it is easy to fall into local minima in the training phase of the network and the convergence speed is slow, which leads to the reduction of classifier recognition rate. To address the above problems, this paper proposes the algorithm of PSO optimized BP neural network (PSO-BPNN) and establishes a classification and recognition model based on this algorithm for college students' entrepreneurship evaluation. The predicted values obtained from the particle swarm optimization neural network model are used to calculate the gray intervals, and the modeling samples are further screened using the gray intervals and the correlation principle, while the hyperspectral particle swarm optimization neural network model of soil organic matter based on the gray intervals is established afterward; and the estimation results are compared and analyzed with those of traditional modeling methods. The results showed that the coefficient of determination of the gray interval-based particle swarm optimization neural network model was 0.8826, and the average relative error was 3.572%, while the coefficient of determination of the particle swarm optimization neural network model was 0.853, and the average relative error was 4.34%; the average relative errors of the BP neural network model, support vector machine model, and multiple linear regression model were 8.79%, 6.717%, and 9.9%, respectively. The average relative errors of the BP neural network model, support vector machine model, and multiple linear regression model are 8.79%, 6.717%, and 9.468%, respectively. In general, the entrepreneurial ability of college students is at a good level (83.42 points), among which the entrepreneurial management ability score (84.30 points) and entrepreneurial spirit (84.16 points) are basically the same, while the entrepreneurial technology ability is relatively low (82.76 points), and the evaluation results are further verified by the double case analysis method. The current problems encountered by university students in entrepreneurship are mainly the lack of practicality, which indicates that universities, industries, and national strategy implementation levels are not sufficiently focused and collaborative in entrepreneurship development to varying degrees.

1. Introduction

With the birth and growth of the Internet, all kinds of related technologies develop at a high speed, the distance between people is drawn closer and closer, and work and life are increasingly inseparable from the Internet. Internet of Things (IoT) technology, as an important branch of Internet technology, combines various information sensing devices with the Internet, completes the transmission and sharing of data through the Internet, and uses computer technology to realize the processing and application of relevant data [1]. This technology extends the information exchange once limited to human to human, to things to people, and to things to things, making the future network more intelligent and realizing the interconnection of all things. With the development of the information technology revolution and the extension of economic globalization, "innovation" has increasingly become a core element in promoting social and

economic development. To meet the trend of the times, the Party and the government have made important strategic plans for building an innovative country, and "innovation" will be the main theme for a long period. Innovation needs to be led by direction, and for this reason, the development paths of "independent innovation", "integrated innovation", and "re-innovation after introduction, digestion, and absorption" have been put forward one after another. To achieve innovation, it is necessary to cultivate and shape many innovators and entrepreneurs who meet the needs of the times [2]. Clarified the relationship between the five major elements of academic evaluation and the subjective and objective factors of creativity training, as well as the mechanism of their influence, through theoretical analysis and empirical research, pointed out that strong chemical industry evaluation's guidance for creativity training is for creativity training, and it is of great significance. How to improve the innovation and entrepreneurial competency of college students has become the focus of research in the academic field. The purpose of this paper is to construct a theoretical model of entrepreneurial competency that is consistent with the characteristics and rules of entrepreneurship development of college students by drawing on existing research results and empirical studies and to empirically evaluate and test the overall level and influencing factors of entrepreneurial competency of college students through the theoretical model, to contribute to the in-depth development of college students' entrepreneurship research and practice [3].

The scientific evaluation of the entrepreneurial competencies of science and technology students and the effective guidance of their entrepreneurial activities are theoretical issues that need to be addressed urgently. Entrepreneurship as a serious socio-economic activity has its logic and laws, and entrepreneurial behavior is in line with the consensus of economics and management, that is, entrepreneurship as a market-oriented behavior requires that entrepreneurs should have a competence structure that meets the needs of entrepreneurship, and it is absolutely difficult to guarantee the smooth development and profitability of entrepreneurs by ignoring the competence elements and to ensure that the country's strategic goal of achieving social stability and economic development through high-quality entrepreneurship of college students can be successfully achieved [4]. It is difficult to ensure that the strategic goal of social stability and economic development can be successfully achieved through the high-quality entrepreneurship of college students. Therefore, the systematic answer to the questions of what is the basic competency structure of entrepreneurship among college students, whether contemporary college students have entrepreneurial competency, the degree of proficiency in their entrepreneurial competency, and how to improve college students' entrepreneurial competency has become a major issue that cannot be avoided in the field of entrepreneurship research. How to effectively identify and construct a competency structure system is not only an inherent requirement of entrepreneurship evaluation for science and technology undergraduates but also an objective need to effectively guide their entrepreneurial activities.

The structural contradiction between university graduates and market demand is increasingly prominent. The talents cultivated by the old training model cannot match the requirements of the economic transformation on the quality and knowledge skills of talents, resulting in the embarrassing situation that enterprises cannot obtain effective talent replenishment in the labor market and college graduates can hardly meet the requirements of job positions. Thus, the participation of college students in entrepreneurship is an important means to enhance their comprehensive quality and improve their social survival ability, and it is also an effective way to drive employment through entrepreneurship and relieve social employment pressure. Entrepreneurial practice allows college students to have a deeper contact with society and understand it, to discover problems, stimulate innovation, solve problems, and create value. For college students, school is only a small classroom for learning cultural knowledge, but society is the big classroom that can bring more learning and growth space for college students. Entrepreneurial practice is like a bridge between school and society, which can greatly accelerate the process of professionalization and socialization of college students and improve their comprehensive ability to adapt to society. To sum up, the era has given innovative, practical, humanistic, and social characteristics to university students' entrepreneurship. It is these characteristics of the times that reflect the importance of guiding and enhancing students' entrepreneurial competence and stimulating and tapping their potential human capital in the context of "building an innovative country, promoting the deepening reform of higher education, and transforming the development connotation of the market economy".

2. Current Status of Research

A model of entrepreneurial process based on the view of entrepreneurship as organizational creation is proposed, which considers that the entrepreneurial process can be divided into two stages: first, the preorganizational venture formation period, in which entrepreneurial activity is mainly expressed as individual entrepreneurial behavior; second, the postorganizational venture operation period, in which entrepreneurial activity is mainly expressed as organizational behavior of the entrepreneurial enterprise [5]. Some scholars, on the other hand, divide the entrepreneurial process into five stages according to the growth process of a new enterprise: first, the proof-of-principle stage, where the main task is to verify the feasibility of the innovative technology; second, the prototype stage, where the product is produced and the prototype organizational structure is formed; third, the model sales stage, where the main task is to improve the feasibility of the product and make the necessary professional division of labor for financial, marketing, and other operational activities; fourth, the start-up stage, where the profitability of the product further increases, the organizational structure expands further, and more product and management issues arise; and fifth is the natural growth stage, where the organization enters natural expansion and the entrepreneur usually considers new

strategic adjustments or enters a new entrepreneurial cycle [6]. Avoid harming the person being evaluated. It is necessary to explore how to cultivate students' creativity in academic evaluation. This is a subject that spans two research fields. Naturally, it is inseparable from the research of creativity cultivation. The existence stage, the main goal is to make business opportunities to form enterprises, where only a small number of enterprises can earn enough cash income to enter the second stage; survival stage, where the goal is to grow the business into the next stage and cannot stagnate in the survival stage for long enough to cause cash flow to dry up; is the success stage, where the organization is taking shape and the business is earning good revenues and needs to adjust its strategy in time to prevent the business from regressing in the face of external changes; the takeover stage, where investors look for opportunities to take the business off the table and voluntary or involuntary changes in management occur; and the resource maturity stage, where the business is operating with abundant resources and effective realization of economies of scale [7].

It is proposed that the entrepreneurial process is a highly dynamic balance between the appropriate allocation of entrepreneurial opportunities, entrepreneurial team, and resources, and entrepreneurial opportunities, resources, and entrepreneurial team are the key components of the entrepreneurial process [8]. In the early stage of entrepreneurship, the key is the exploration and selection of opportunities, and the decision of the entrepreneurial team is focused on the rapid integration of resources to seize the entrepreneurial opportunities; as the new enterprise is founded and grown, resources become more abundant, and the enterprise faces a more complex competitive environment and market environment, the decision of the entrepreneurial team is focused on the rational allocation of resources to improve the efficiency of resource use, and the construction of a standardized management system to resist external competition and uncertainty [9]. A theoretical model of the entrepreneurial process based on the interaction between the entrepreneur and the new enterprise is proposed, arguing that the entrepreneur and the new enterprise are the key components of the entrepreneurial process and that the entrepreneurial process is essentially a process of close interaction between the entrepreneur and the new enterprise under the action of the external environment [10]. The only way to make entrepreneurship research more systematic based on diversification is to gradually build a framework for entrepreneurship research independent of other fields, and then provide a more comprehensive longitudinal research on entrepreneurial competency, entrepreneurial ability, and another individual level [11].

The entrepreneurial process concretely expressed as the act and activity of achieving stage equilibrium in a constant dynamic change. In the entrepreneurial process, there is never absolute stability and equilibrium, and any enterprise and organization are in dynamic change to seek temporary unity and coordination of the elements of entrepreneurship. In different stages of development, entrepreneurs need to analyze the problems and risks that may arise in the process of development from the perspective of dynamic balance and make changes and deployment of their capabilities and organizational resources according to the time and the actual situation, to achieve sustainable development of the enterprise.

3. Adaptive Mutation of the Internet of Things PSO-BPNN Algorithm-Assisted Evaluation of College Student Entrepreneurship Analysis

3.1. IoT Adaptive Mutation PSO-BPNN Algorithm Design. Random access (RA) refers to the process in which the UE or terminal sends an access request to the access control system, and the system allocates access resources such as access channels and time-frequency resource blocks after a certain algorithm and feeds this information back to the terminal, while the data transmission process generally arranged after obtaining the allocated access resources [12]. Random-access methods can be divided into two categories based on the different resource allocation methods: competition-based random-access methods and noncompetition-based random-access methods. The biggest difference between competition-based and noncompetition-based random-access methods is the different allocation of the preamble codes, competition-based random-access method by the UE side to initiate the application of the preamble codes, and the UEs do not know whether the application of the preamble codes is duplicated, so there is a problem of collision, while noncompetition-based random-access method by the access control side of the unified allocation of the preamble codes, so there is no possibility of conflict. The access process also does not require the steps of applying for the leading code and competition resolution.

In the daily access process, uniformly arriving applications often do not cause congestion, and only the influx of many access applications in a short period will lead to severe congestion in the system. Therefore, to simplify the user access flow model, some literature only considers the impact of a one-time large number of arriving access applications on the system and only considers the one-time arrival of N MTC devices at a certain moment in the simulation, without considering whether the subsequent moments. The simulation only considers the arrival of N MTC devices at one time and does not consider whether there are further arrivals of new access requests at subsequent times [13]. This method is less used in the literature but is often used for qualitative analysis of access algorithms, such as the impact of different leading code values on the system access performance. This type of method is persistent over some time with a certain number of access applications arriving, and the specific number of arrivals per time slot can conform to different distribution rates. This type of access flow model is more in line with the actual access in use and is more dynamic and continuous than the first type of access flow model. Depending on the arrival rate of each time slot, this type of access flow model can also be divided into several categories, as shown in Figure 1. At the same time, it also pays attention to the improvement of the comprehensive quality and practical ability of innovation and entrepreneurship,
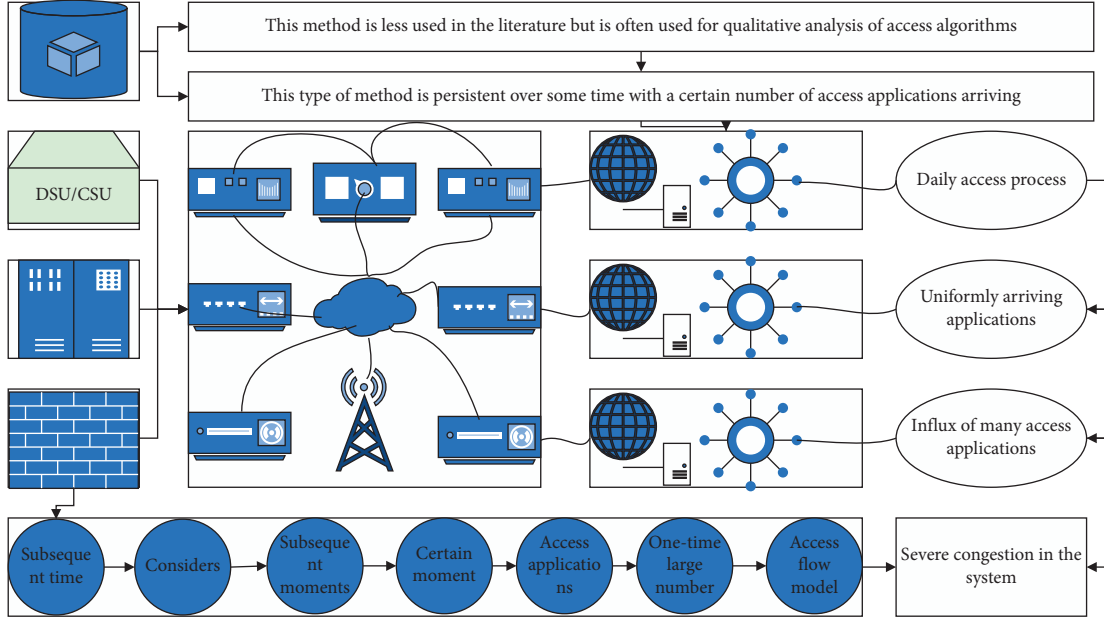
Figure 1: IoT adaptation mutation framework.

especially the expansion of self-entrepreneurship awareness and innovative operation ability, so that students can find and solve problems independently, and then put forward their own new views and opinions, for the life of students The continuous development of learning and the individual has laid a solid foundation.

In the IoT mass access scenario, for periodic services such as the regular reporting of sensor data is suitable for uniformly distributed or Poisson distributed access flow model, while for bursty services such as alarm-like services is suitable for beta distributed access flow model, IoT as an emerging application in the communication industry, driven by big data and artificial intelligence, the market scale will further expand, with technological progress [14]. With technological advancement, government support, and increasingly improved industry standards, China's IoT industry will continue its good development momentum and provide new impetus for sustained and stable economic growth. The wave of expansion of mobile Internet to the Internet of everything will create larger market space and industrial opportunities compared to the Internet [9]. IoT has become an important technical support for many high-tech enterprises in China, such as cloud computing, mobile payment system, bicycle-sharing, and other IoT-related industrial chains, and have achieved great success in the Chinese market, bringing good economic benefits while also bringing great convenience to our travel and work.

Network sizing estimation is usually a comprehensive analysis of both coverage and capacity to determine the size, i.e., the number of sites, required for network planning in each area. On the one hand, the coverage estimation combines the link budget with the propagation model to calculate the coverage radius of the base stations to find the number of base stations required to cover the planning area; on the other hand, the capacity estimation calculates the number of base stations required to meet all the service demands by processing various actual services into some virtual equivalent services based on the resources provided by individual base stations. The larger of the coverage estimate and the capacity estimate are used as the number of sites for the size estimate.

$$
\begin{aligned}
R_{\text{SBS}} &= \left\{ d_{sll} | \text{PL}\left(d_{su}\right) = \chi_{\text{PL}}^2 \right\}, \\
N_{\text{cap}} &= \sum_{k=1}^{N_{\text{region}}} N_{\text{cap}}^k \chi_{\text{PL}}^2.
\end{aligned}
\tag{1}
$$

Among the various types of artificial neural networks, multilayer feedforward networks can not only solve the heterogeneous problem and the arbitrary function approximation problem but also the classifier constructed based on this network has a strong classification recognition ability. BPNN is a typical feedforward neural network, but it has slow convergence and weak generalization ability. Many practical problems have many local minima due to the complex multidimensional nature of the problem itself, and because the BP algorithm itself uses the gradient descent method to find the optimal value will also lead to local minima, resulting in the overall design of the network is more likely to fall into local minima. To solve the above problems of convergence and local minima in BPNN, this paper proposes to optimize the training neural network with the help of a particle swarm algorithm.

The PSO algorithm optimizes the BPNN by replacing the gradient descent method to correct the network weights and thresholds. The key point to consider in this process is to construct a mapping between the neural network weights and thresholds and the particle dimensions of the PSO algorithm. The key to the training of the BPNN is the continuous correction of the weights and thresholds, while the PSO algorithm is the iteration of the individual particle positions and velocities [15]. Therefore, when the two

algorithms are combined, the weights and thresholds of the BP algorithm correspond to the particle positions of the PSO algorithm, and the total number of weights and thresholds is the dimensionality of the particle population. The fitness function of the particles is generally chosen as the objective function, so the fitness function of the particles is determined as MSE according to the end condition of the BP algorithm as the reference minimum mean square error (MSE).

$$\text{MSE} = \frac{1}{P} \sum_{p=1}^{P} \sum_{k=1}^{m} \left(T_k^p + Y_{k+1}^p \right)^2. \quad (2)$$

In this paper, the principle of the PSO-BPNN algorithm is to optimize the parameters of BPNN by particle swarm algorithm training, mainly optimizing the weights and thresholds, which can effectively solve the problem of BPNN falling into local miniaturization and slow convergence, and improve the training efficiency of the network. PSO algorithm in training optimization of BPNN, the weights, and thresholds of BPNN is derived from the position vector of particles. In the search process of the particles, the minimum is reached using the following equation:

$$\text{Fitness}_i = \frac{1}{N} \sum_{i=1}^{P} \sum_{j=1}^{m} \left(W_{ij}^d - W_{ji}^p \right)^2. \quad (3)$$

Innovation and entrepreneurship education is oriented to all students, emphasizing not only the cultivation of basic literacy, innovation spirit, entrepreneurial consciousness, and entrepreneurial ability but also the improvement of comprehensive quality and practical ability of innovation and entrepreneurship, especially the expansion of self-entrepreneurial consciousness and innovative operation ability, so that students can independently identify problems and solve them, and then put forward their new views and opinions, which lay a solid foundation for students' lifelong learning and individual. The weights and thresholds of the BP algorithm correspond to the particle positions of the PSO algorithm, and the total number of weights and thresholds is the dimension of the particle swarm. This provides students with a solid foundation for lifelong learning and sustainable development [16]. The purpose of innovation and entrepreneurship education is not to let every student start a business, but the core is to cultivate students' innovation spirit, entrepreneurial consciousness, and entrepreneurial ability, focus on the cultivation of students' independent learning and practical ability, closely combine the cultivation of talents with scientific research and the needs of society, and change from the traditional classroom focus on knowledge transfer to the emphasis on ability and comprehensive quality cultivation, which can effectively improve the quality of talent cultivation in colleges and universities, as shown in Figure 2.

Educational evaluation should be as impartial and objective as possible to make the results of the evaluation have a greater reference value. In the development of educational evaluation activities, objective facts should be used as the
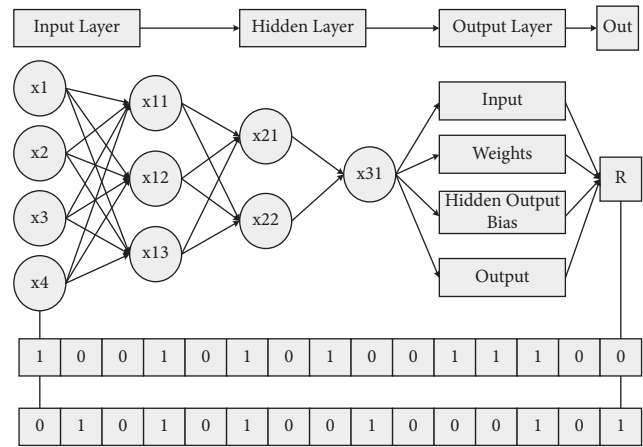
Figure 2: PSO-BPNN algorithm framework.

basis, and the reference standard for evaluation should be used as the yardstick, eliminating as far as possible the interference of the evaluator's subjective factors and external objective factors in the evaluation. Both the methods and means of evaluation and the standards of evaluation should be in line with the norms of ethics and morality. Once the evaluation criteria have been established, they should be strictly enforced from the beginning to the end. In addition, access to evaluation information should be justified and the use of evaluation results should be legal to avoid harming those being evaluated. To explore how to cultivate students' creativity in academic assessment, which is a topic that spans two research fields, it is natural that the research on creativity cultivation is inseparable. There are many research results, both theoretical and practical, on creativity education and creativity cultivation. Analysis of the components of creativity and its characteristics can provide a more scientific theoretical basis for carrying out academic assessment reform.

3.2. Design of Entrepreneurship Evaluation Model for College Students in Higher Education. The innovation of this research lies in: first, the innovation of research perspective. There have been studies on the cultivation of creativity, mainly from the perspective of optimizing curriculum structure, integrating learning resources, enriching learning experience, reforming teaching methods and approaches to talk about the cultivation of creativity, but this study starts from academic assessment to discuss the issue of creativity cultivation, which is an innovation in terms of research perspective. Second, it is an innovation in research content. Using the modern educational evaluation theory and the theory of creativity education, this study provides a systematic and comprehensive discussion on the promotion of creativity cultivation by academic evaluation in colleges and universities, clarifies the correlation between the five major elements of academic evaluation and the subjective and objective factors of creativity cultivation, as well as the mechanism of their influence [17]. Through theoretical analysis and empirical research, it is pointed out that strengthening the orientation of academic evaluation on

creativity cultivation is of great significance to creativity cultivation, summarizing the academic evaluation reform strategy of "five guides," constructing a framework system for academic evaluation of college students, and proposing specific operational suggestions for the reform of evaluation contents and evaluation methods of different types of coursework. In the analysis of the problem by AHP, the relevant influencing factors and solutions are listed according to the overall objective of the problem, and a hierarchical model is constructed. The constructed hierarchical model generally consists of three levels, namely, the objective level, the criterion level, and the solution level. The criterion layer, also known as the intermediate layer, represents the factors that need to be considered and the guidelines for decision making to achieve the overall objective and is the intermediate link to achieve the overall objective, as shown in Figure 3.

The elements of the evaluation of graduate students' innovation and entrepreneurship education initially formulated in this study are mainly analyzed in terms of top-level design, teachers, students, and educational environment and forms. After consulting experts and synthesizing their opinions, the top-level design mainly designed two evaluation dimensions, namely, the incentive support policies related to graduate students' innovation and entrepreneurship and the evaluation of universities' attention to graduate students' innovation and entrepreneurship education. The study explores the influencing factors of postgraduate innovation and entrepreneurship and then designs a framework of postgraduate innovation and entrepreneurship education evaluation system with high realistic guidance value. Moreover, the classifier based on this network has strong classification and recognition capabilities. Therefore, exploring the optimization problem of feedforward neural networks has practical significance for the research in many fields.

The support of graduate student innovation and entrepreneurship education and the teaching methods of graduate student innovation and entrepreneurship education, coded as DA1, DA2, and DA3 in order; the measurable indicators of faculty provision are the faculty provision of graduate student innovation education and the faculty provision of graduate student entrepreneurship education, coded as IF1 and IF2 in order; and the measurable indicators of implementation methods are the practical forms of graduate student innovation education and the practical forms of graduate student entrepreneurship education. Based on the questionnaire survey, to further explore the relationship between the influencing factors, a theoretical model on graduate innovation and entrepreneurship education was constructed with the help of structural equation modeling, as shown in Figure 4.

After running the standardized estimates calculation, the results of the unstandardized path coefficients of the model show that all the path coefficients except the 11 path coefficient reference indicators reach a significant level of 0.05, the measurement errors are positive, the absolute values of the critical ratios are greater than 1.96, and the significance probability values are less than 0.001, indicating that all the

path coefficients are significant and the constructed [18]. The structural equation theoretical model constructed is not problematic. The overall index of graduate students' information literacy is 52.4, which is consistent with the current situation of graduate students' innovation and entrepreneurship education, indicating that the indexes of the influencing factors of graduate students' innovation and entrepreneurship education selected in this study are more effective. However, the value of the total index is only slightly greater than 50, which is not a high index, reflecting that the current situation of the development of innovation and entrepreneurship education for graduate students is not optimistic. The existing evaluation structure of entrepreneurial competency of college students is mainly derived from the evaluation model of entrepreneurial competency of Western entrepreneurs, which is thinly localized on the one hand; and on the other hand, it is difficult to apply the entrepreneurial competency model of general entrepreneurs to college students' entrepreneurs.

Moreover, the current entrepreneurial competency research not highly differentiated from the competency research, which makes it difficult to reflect the theoretical depth of entrepreneurial competency research. Such a research pattern is not conducive to the scientific development of entrepreneurship among college students in science and technology, nor is it conducive to the scientific decision making of management agencies [19]. To determine the scale required for network planning in a given area, that is, the number of sites. Coverage estimation should combine the link budget with the propagation model, calculate the coverage radius of the base station, and find the number of base stations required to cover the planned area. This study precisely seeks to make a breakthrough in this field, deepen theoretically the understanding of entrepreneurial competence, dig deeper into the influencing factors of entrepreneurial competence and countermeasures of cultivation from the theory of college students' entrepreneurship, and strengthen the theoretical explanatory power of evaluation research with vivid cases and an appropriate amount of quantitative research methods to form a theoretical research structure combining depth and breadth, as shown in Figure 5.

This study takes the evaluation of the ability of science and technology students' entrepreneurship as the research theme, studies the current operation mechanism of entrepreneurship education, constructs a system of ability characteristics based on the comprehensive use of scientific methods, relies on field research and quantitative analysis to process, and organize the system, finally constructs an evaluation index system and mathematical model, and verifies it through empirical methods. The main research methods used include literature research method, questionnaire survey method, statistical analysis method, case study method, expert interview method, etc. Each method is not applied independently, but fully cooperated, repeatedly used, and mutually corroborated. Through this study, we can deepen the theoretical depth of college students' entrepreneurial ability evaluation research, expand the field of college students' entrepreneurial ability evaluation research to the
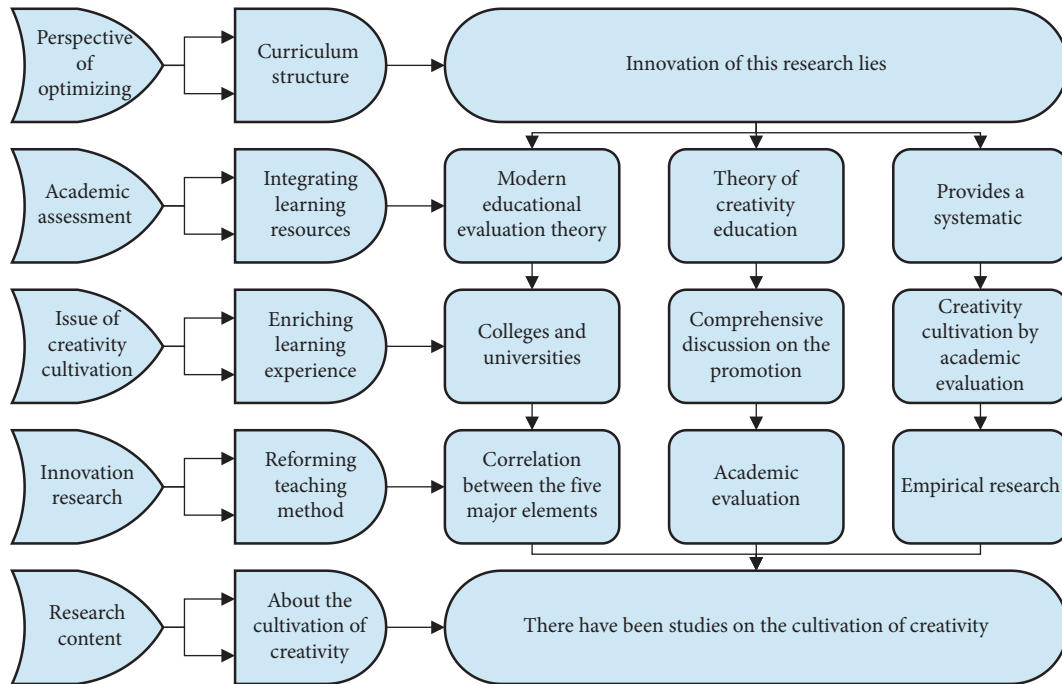
FIGURE 3: Evaluation hierarchy model.

FIGURE 4: Structural distribution of the questionnaire sample.

Guiding and encouraging members to let go of their psychological burdens, delegating a certain degree and scope of authority, and focusing on nurturing team members' responsibility, sense of mission, and trust are the main focuses of team leaders. In terms of mechanism design for implementing the path, a reasonable and effective system of incentives, constraints, and penalties should be gradually established and improved.

4. Analysis of Results

4.1. IoT Adaptation Mutation PSO-BPNN Algorithm Results. In the initialization of the particle swarm, the particle swarm size $M = 60$, the particle swarm size is too small will not be obvious for the optimization of the later BPNN, when M is less than 50, PSO optimization of BPNN will still exist easy to fall into the condition of local minima, when M is greater than 50 then PSO optimization of BPNN is obvious. The scale of the market will be further expanded. With technological progress, government support, and increasingly perfect industry standards, the Internet of Things industry will continue to maintain a good momentum of development and provide new impetus for sustained and stable economic growth. The sum of the number of weights and thresholds in BPNN is 305, so this particle population has dimension $D = 305$, and usually the fitness function is the training error function MSE of the BPNN. The number of iterations is defined as 200, the learning factors c, and care both 2, and the inertia weight is 0.5. The output of the PSO algorithm is stored as weights and thresholds to facilitate the training of the BPNN in the next step. The PSO algorithm is selected instead of the gradient descent method to correct the weights and thresholds and start training the BPNN so that the training accuracy reaches 0.00001 or the defined

group of science and technology college students, and propose a model of evaluation of college students' entrepreneurial ability in science and technology with the characteristics of the times, reflecting the spirit of China and in line with the reality of science and technology college students and verify it [20]. The biggest potential problem at this stage is that some team members may have palpitations about the difficulties of the turbulent period and cannot express their opinions and insights effectively and sincerely.

Figure 5: Theoretical path diagram of the structural equation model.

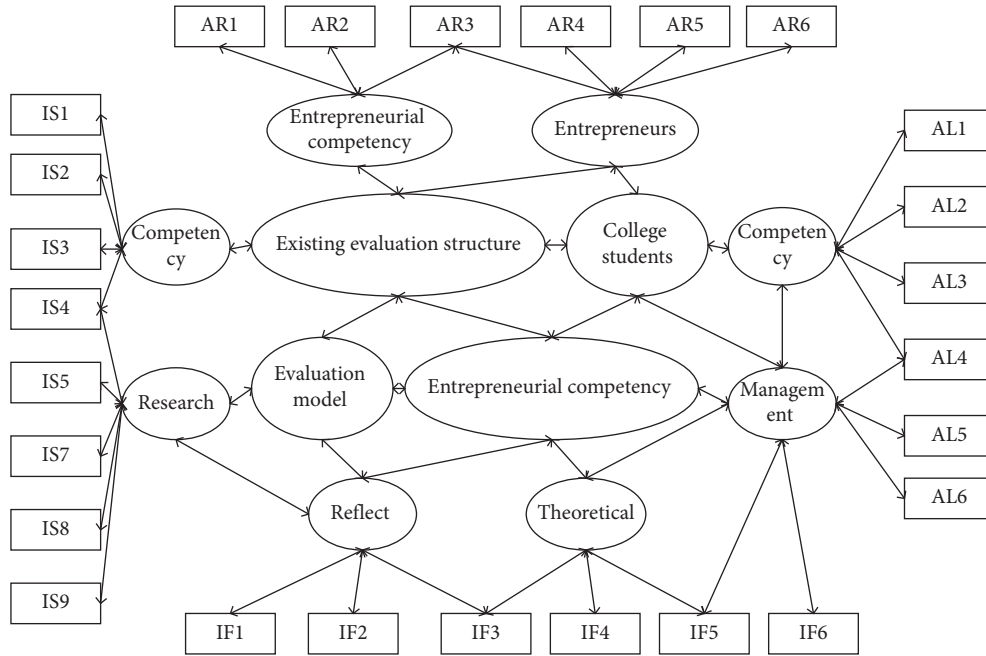number of training times 1000 times ends with a training learning rate of 0.1. When the net training is learned, the feature vectors of any 30 groups of samples for each Chinese character are mixed for training, the total number of training sample images is 960, and the final training samples are built for the next test recognition. The PSO algorithm optimizes the iterative fitness values for BPNN, as shown in Figure 6.

Figure 6 shows that the performance is only improved by about 3% compared to randomly selecting a single relay for transmission, which is much lower than the QL-RSA algorithm under DF protocol. It can be seen from Figure 7 that the performance of the QL-RSA algorithm is not much optimized with random selection as the total transmit power increases. Specifically, the arrival amount of each time slot can conform to different distribution rates. This type of access flow model is more in line with the actual access situation and is more dynamic and continuous than the first type of access flow model. This is because the amplified forwarding protocol amplifies the transmit power of the useful signal at the second hop in addition to the amplified noise power, thus resulting in the received signal-to-noise ratio at the destination of each node is not very different, and the received signal-to-noise ratio is used as a reward for Q-learning, then the difference in the Q-value of each node is not too big, and the Softmax selection strategy is precisely based on the corresponding Q-value of each node. The Softmax selection strategy is based on the corresponding Q-value of each node to make a decision. This is the reason the performance of the QL-RSA algorithm is not much optimized than that of R-RSA.

This is because when the signal-to-noise ratio required at the receiving end is not high, anyone relay node can meet the requirements at the receiving end, so the number of nodes for the three algorithms does not differ much. When the total number of relays is larger, the number of relays
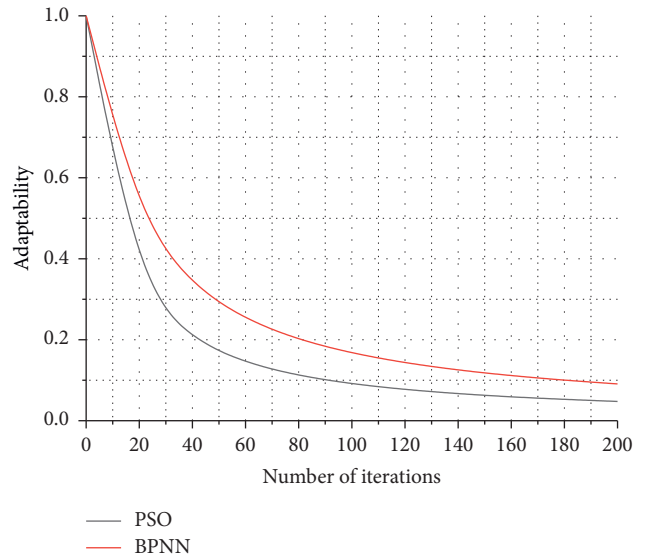
Figure 6: Iterative adaptation degree value graph.

required for cooperative transmission is smaller, because the more densely distributed the relays are, the better the combination of the selected set of relays can be, so that a larger signal-to-noise ratio can be achieved at the receiving end, which in turn means that fewer relays can be used to achieve the upper bound on the signal-to-noise ratio at the receiving end. The optimal single-relay selection and the multirelay selection are discussed separately for the two forwarding protocols. The data transmission process is generally arranged after the allocated access resources are obtained. Random-access methods can be divided into contention-based random-access methods and noncompetition-based random-access methods according to different resource allocation methods. As a comparison of simulation
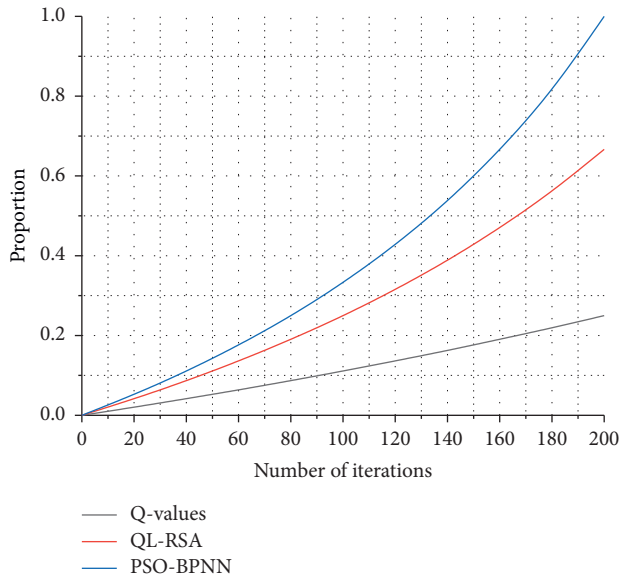
FIGURE 7: Percentage of ideal throughput accounted for by different algorithms.

results, this paper also simulates the results of the R-RSA algorithm and analyzes the performance of the QL-RSA algorithm from different perspectives of average throughput, the probability distribution of selected nodes, and accounting for the ideal average throughput, the simulation results show that the performance of QL-RSA is obvious compared to R-RSA when DF forwarding protocol is used, but the effect is not so obvious for AF forwarding protocol.

4.2. Results of the Evaluation of Entrepreneurship among College Students in Higher Education. In general, the entrepreneurial ability of science and technology college students is at a good level, among which entrepreneurial management ability and entrepreneurship are equal, and entrepreneurial technology ability is relatively low. The entrepreneurial technology ability of science and technology college students is the center of gravity in their entrepreneurial ability structure, but it is also a relative shortboard that still has room for improvement. The above phenomenon can be tentatively explained by referring to Timmons' entrepreneurial process theory. While solving its own employment, it creates more opportunities and jobs for the society. The participation of college students in entrepreneurship is an important means to enhance their own comprehensive quality and improve their social viability, and it is also an effective way to promote employment through entrepreneurship and relieve social employment pressure. In the early stage of entrepreneurship, entrepreneurs are required to have deep and good professional abilities. For science and technology undergraduates, the level of technical ability can largely influence their quality of exploring and grasping opportunities, while the improvement of technical ability is quite difficult. The influence of scientific and technological ability on entrepreneurial management ability and entrepreneurial

spirits, such as resource application and team management, is weak, and such management ability and entrepreneurial spirit are less required in the early stage of entrepreneurship, and college students can basically reach the initial response after a few years of systematic higher education, so the evaluation is good. However, it should be noted that in the midstage of entrepreneurship, the requirements for resource application and management skills are substantially increased, and the reorganization of resources, the recreation of opportunities, and the operation of teams are raised to a strategic level. The subjects observed in this study are mainly university students, which are in line with the competence performance of subjects in the early or even pre-entrepreneurial stage of entrepreneurship, as shown in Figure 8.

Teaching induction requires teachers to accept and adopt new teaching methods. In entrepreneurship teaching activities for science and technology college students, teachers can use task-driven and typical demonstration teaching methods to induce students to spontaneously generate entrepreneurial views and consolidate entrepreneurial enthusiasm to develop entrepreneurial ability. In the task-driven segment, teachers set specific tasks with technical points and difficulties of technology-based entrepreneurial activities, ask students to complete the tasks in groups and give comments and scores according to the process and results of the groups' completion of the tasks, which induces students' enthusiasm for competition and teamwork, and gains professional knowledge and experience in the process of completing the tasks.

In the typical demonstration, session teachers introduce skilled entrepreneurs or R&D leaders in enterprises into the school to teach the more practical and operational aspects of the teaching. These entrepreneurs are successful models of real entrepreneurship, and interactive exchanges with entrepreneurs can induce students' entrepreneurial enthusiasm perceptually, effectively stimulate students' sense of imitation, consolidate students' enduring interest in entrepreneurship, and enhance the bonding of entrepreneurship education. To enhance the effectiveness of teaching induction, we can consider organizational restructuring by establishing a special entrepreneurship center or entrepreneurship college to cooperate with the technological innovation centers, product development centers, and employment counseling agencies that are currently prevalent in universities, to bring into play the entrepreneurship education and counseling functions of these organizations. It has very important theoretical value and innovative significance for enriching and perfecting the evaluation theory of graduate students' innovation and entrepreneurship education. For example, the Entrepreneurship Center is responsible for providing resources for research and development of entrepreneurship courses, strengthening links with engineering and technology-related industries and academia through cooperation with science and technology transformation departments, and publishing horizontal research topics on innovation and entrepreneurship. It can also cooperate with on-campus
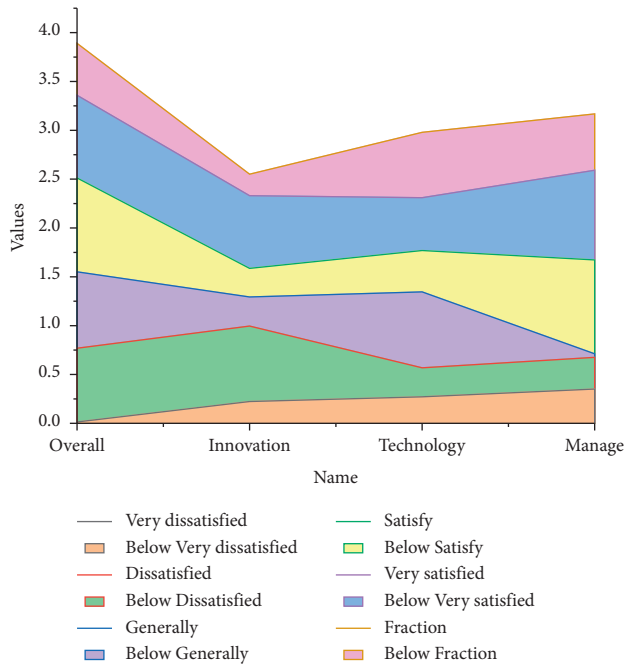
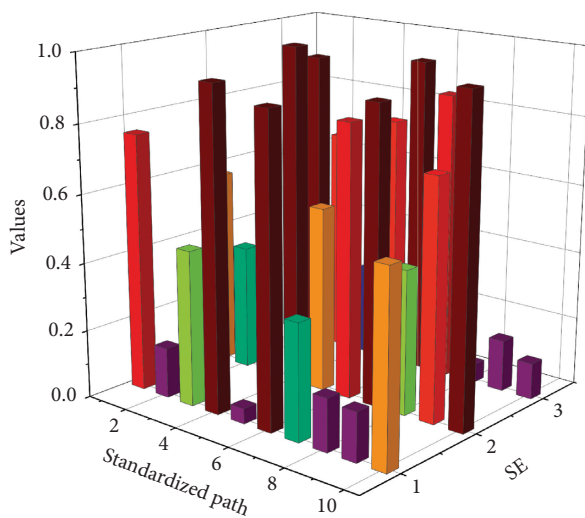FIGURE 8: College students' entrepreneurship scores.

FIGURE 9: Path coefficients of the modified entrepreneurial opportunity capability structure model.

employment counseling agencies to provide students with entrepreneurial counseling and practical platform services in terms of industry and career development prospects to enhance students' confidence in entrepreneurship, as shown in Figure 9.

To further illustrate the relationship between the factors of entrepreneurship development resources and entrepreneurial competencies, a structural equation model was conducted separately in this chapter to test the relationship between entrepreneurship development resources and entrepreneurial generic competencies, entrepreneurial professional competencies, and entrepreneurial opportunity competencies, and the results showed that the effects of the factors on entrepreneurship generic competencies and

entrepreneurial professional competencies were social network resources, practical training resources, education and teaching resources, and policy and cultural resources, respectively. How to improve the systematic response to practical problems such as college students' entrepreneurial ability has become a major issue that cannot be avoided in the field of entrepreneurship research. The results showed that the effects of the factors on entrepreneurial general ability and entrepreneurial professional ability were social network resources, practical training resources, education and teaching resources, and policy and cultural resources, respectively, and the effects of the factors on entrepreneurial opportunity ability were social network resources, education and teaching resources, and practical training resources, respectively, and policy and cultural resources did not have an effective effect on entrepreneurial opportunity ability.

5. Conclusion

This paper outlines the principle, structure, and implementation process of BPNN and its properties. The PSO-BPNN algorithm is proposed to optimize the BPNN with the PSO algorithm for the defects of slow convergence and easy to fall into the local minima of BPNN. And the prediction feasibility of the optimized algorithm is verified by function simulation fitting, and the optimized PSO-BPNN algorithm has the advantages of high prediction accuracy, fast convergence, and fewer training times than the original BPNN algorithm. Based on the predicted value of the application volume and based on the principle of the maximum number of successful next time slot accesses, the leading code grouping threshold and the ACB control parameters are adjusted, and then the newly arrived service types are identified according to the QoS requirements and the size of the data volume, and the previously derived parameters are used, the access method of time slot Aloha is adopted for the services with insensitive delay and small effective data volume and the cultivation of entrepreneurial ability of college students. The purpose is to build a set of competency theoretical models that conform to the characteristics and laws of college students' entrepreneurial development by referring to existing research results and empirical research. This study shows that the effect of the factors on entrepreneurial general ability and entrepreneurial professional ability is social network resources, practical training resources, education and teaching resources, and policy and cultural resources, respectively; the effect of the factors on entrepreneurial opportunity ability is social network resources, education and teaching resources, and practical training resources, respectively.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] C. Brentnall, I. D. Rodríguez, and N. Culkin, "The contribution of realist evaluation to critical analysis of the effectiveness of entrepreneurship education competitions," *Industry and Higher Education*, vol. 32, no. 6, pp. 405–417, 2018.

[2] H. Jiang, W. Xiong, and Y. Cao, "Research on the mechanism of entrepreneurial education quality, entrepreneurial self-efficacy and entrepreneurial intention in social sciences, engineering and science education," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 7, pp. 3709–3721, 2017.

[3] F. Wright, "How do entrepreneurs obtain financing? An evaluation of available options and how they fit into the current entrepreneurial ecosystem," *Journal of Business & Finance Librarianship*, vol. 22, no. 3-4, pp. 190–200, 2017.

[4] A. Yulastri and H. Hidayat, "Developing an entrepreneurship module by using product-based learning approach in vocational education," *International Journal of Environmental & Science Education*, vol. 12, no. 5, pp. 1097–1109, 2017.

[5] Y. Y. Ding, "The constraints of innovation and entrepreneurship education for university students," *Journal of Interdisciplinary Mathematics*, vol. 20, no. 6-7, pp. 1431–1434, 2017.

[6] F. Liñán, F. Ceresia, and A. Bernal, "Who intends to enroll in entrepreneurship education? Entrepreneurial self-identity as a precursor[J]," *Entrepreneurship Education and Pedagogy*, vol. 1, no. 3, pp. 222–242, 2018.

[7] H. Ni and Y. Ye, "Entrepreneurship education matters: exploring secondary vocational school students' entrepreneurial intention in China," *The Asia-Pacific Education Researcher*, vol. 27, no. 5, pp. 409–418, 2018.

[8] A. S. Alshebami, I. Al-Jubari, I. Y. Alyoussef, and M. Raza, "Entrepreneurial education as a predicator of community college of Abqaiq students' entrepreneurial intention," *Management Science Letters*, vol. 10, no. 15, pp. 3605–3612, 2020.

[9] C. Bandera, R. Collins, and K. Passerini, "Risky business: experiential learning, information and communications technology, and risk-taking attitudes in entrepreneurship education," *International Journal of Management in Education*, vol. 16, no. 2, pp. 224–238, 2018.

[10] D. K. Hsu, K. Burmeister-Lamp, S. A. Simmons, M.-D. Foo, M. C. Hong, and J. D. Pipes, ""I know I can, but I don't fit": perceived fit, self-efficacy, and entrepreneurial intention," *Journal of Business Venturing*, vol. 34, no. 2, pp. 311–326, 2019.

[11] T. Anjum, S. R. Ramzani, M. Farrukh, V. Raju, N. Nazar, and I. A. Shahzad, "Entrepreneurial intentions of Pakistani students: the role of entrepreneurial education, creativity disposition, invention passion & passion for founding," *Journal of Management Research*, vol. 10, no. 3, pp. 76–100, 2018.

[12] E. Noyes, "Teaching entrepreneurial action through prototyping: the prototype-it challenge," *Entrepreneurship Education and Pedagogy*, vol. 1, no. 1, pp. 118–134, 2018.

[13] K. Nuringsih, M. N. Nuryasman, and R. A. IwanPrasodjo, "Sustainable entrepreneurial intention: the perceived of triple bottom line among female students," *Jurnal Manajemen*, vol. 23, no. 2, pp. 168–190, 2019.

[14] J. L. Burnette, J. M. Pollack, R. B. Forsyth et al., "A growth mindset intervention: enhancing students' entrepreneurial self-efficacy and career development," *Entrepreneurship: Theory and Practice*, vol. 44, no. 5, pp. 878–908, 2020.

[15] M. Naushad, "A study on the antecedents of entrepreneurial intentions among Saudi students," *Entrepreneurship and Sustainability Issues*, vol. 5, no. 3, pp. 600–617, 2018.

[16] H. V. Mukesh, A. S. Rao, and K. Rajasekharan Pillai, "Entrepreneurial potential and higher education system in India," *Journal of Entrepreneurship*, vol. 27, no. 2, pp. 258–276, 2018.

[17] R. Badri and N. Hachicha, "Entrepreneurship education and its impact on students' intention to start up: a sample case study of students from two Tunisian universities," *International Journal of Management in Education*, vol. 17, no. 2, pp. 182–190, 2019.

[18] C. Palmer, U. Fasbender, S. Kraus, S. Birkner, and N. Kailer, "A chip off the old block? The role of dominance and parental entrepreneurship for entrepreneurial intention," *Review of Managerial Science*, vol. 15, no. 2, pp. 287–307, 2021.

[19] X. Qiao and J. H. Huang, "Effect of college students' entrepreneurial self-efficacy on entrepreneurial intention: career adaptability as a mediating variable," *International Journal of Educational Methodology*, vol. 5, no. 3, pp. 305–313, 2019.

[20] M. Rodriguez, S. Boyer, D. Fleming, and S. Cohen, "Managing the next generation of sales, gen Z/millennial cusp: an exploration of grit, entrepreneurship, and loyalty," *Journal of Business-To-Business Marketing*, vol. 26, no. 1, pp. 43–55, 2019.

Research Article

Research and Implementation of Mobile Internet Management Optimization and Intelligent Information System Based on Smart Decision

Yanqing Han [ID],[1] **Yuyan Lei** [ID],[2] **Zimin Bao** [ID],[1] **and Qingyuan Zhou** [ID][3]

[1]*Xi'an Jiaotong University City College, Xi'an, China*
[2]*College of Economics and Management, Xi'an University of Posts & Telecommunications, Xi'an, China*
[3]*School of Economics and Management, Changzhou Vocational Institute of Mechatronic Technology, Changzhou, China*

Correspondence should be addressed to Yuyan Lei; leiyuyan@xupt.edu.cn and Qingyuan Zhou; zqy2131@czimt.edu.cn

The way by which artificial intelligence is implemented is similar to the thinking process of the human brain. People obtain information about external conditions through five senses, namely, vision, hearing, smell, taste, and touch, and, through the further processing of the brain, it forms meaningful decision-making elements. Then, through the process of analysis and reasoning, further decisions are made. In the information age, the application of intelligent management information systems in various fields has promoted the modernization and intelligence of social development. From the perspective of intelligent decision-making, this paper analyzes the requirements of intelligent information systems and designs an intelligent information system based on mobile Internet management optimization, including system management optimization, and proposes an environment-based layer, network transport layer, and the three-tier system architecture of the smart service application layer. Finally, this paper considers the problem of data fusion after system expansion. According to the existing fuzzy fusion algorithm, a weight-based fuzzy fusion algorithm is proposed. The simulation analysis shows that the algorithm can be effectively applied in intelligent information systems.

1. Introduction

Artificial intelligence (AI) is an important branch of computer science. It attempts to understand the essence of intelligence and produce a new intelligent machine that can respond in a similar way to human intelligence. The research in this field includes robot, language recognition, image recognition, natural language processing, and expert system. Artificial intelligence can simulate the information process of human consciousness and thinking. Artificial intelligence is not human intelligence, but it can think like people and may exceed human intelligence. It is also considered as one of the three cutting-edge technologies in the 21st century. In the past few years, some computer systems with artificial intelligence have been established to control spacecraft and underwater robots [1]. Through the program, people can make some thinking reasoning, so that they have certain advanced intelligences such as environmental adaptation, automatic learning, and automatic decision-making [2]. In addition to the application of artificial intelligence to sensory simulation, a more important application is to simulate the thinking and analysis process of the human brain, namely, game and logical reasoning, the application of information sensing, and processing. In games or simulation systems, the application of such artificial intelligence is extraordinarily rich. Some techniques, applied in the chess program, such as looking forward a few steps and breaking down complex problems into some easy subproblems, are developed. It evolved into the basic technology of artificial intelligence such as search and problem summarization. At present, the project technology is developing rapidly and amazingly [3]. The next generation of artificial intelligence system will

affect our life more widely. Artificial intelligence will make more critical and personalized decisions for human beings through interaction with the environment.

The research goal of "3S Smart Service System" [4] is to build a universal access system based on technologies such as Internet of Things, cloud computing, and big data, which can accept intelligent industry, intelligent agriculture, intelligent logistics, intelligent transportation, smart grid, intelligent environmental protection, intelligent security, intelligent medical treatment and smart home, etc. Different types of IoT specific applications [5] provide common solutions for different application scenarios of IoT technology and achieve unified management and control. Considering that the system will be built into an intelligent system with massive data, scientific control, intelligent services, etc., the system is particularly important for the storage, management, processing, and transmission of data. The secure storage, specification management, intelligent processing, and reliable transmission of service environment data will provide reliable guarantee for data volume and data accuracy for system computing and decision-making.

Aiming at the problems existing in intelligent products or systems, this paper intends to adopt a support system model based on intelligent decision-making to design an intelligent information system based on Internet management optimization. The research focuses on intelligent information systems to integrate different intelligent products into the same system and, through big data and cloud computing technology, provide users with truly intelligent, personalized, automated full-service, without the need to control the system through user behavior. Considering the massive data characteristics of the system expansion, according to the existing fuzzy fusion algorithm, a fuzzy fusion algorithm based on weight is designed to provide more accurate and reliable raw data for the decision module of the system.

The remains of this paper are organized into four sections. Section 2 contains related works of our research field. In Section 3, we give the design of system model. Section 4 covers on the experimental and results. In the last section, we draw the discussion and conclusion over our research.

2. Related Work

Decision support system (DSS) is a computer application system that assists decision-makers to make semistructured or unstructured decisions by human-computer interaction through data, model, and knowledge [7]. It is an advanced information management system produced by the development of management information system (MIS) to a higher level. It provides an environment for decision- makers to analyze problems, establish models, simulate decision-making processes and schemes, and call various information resources and analysis tools to help decision-makers improve decision-making level and quality. Because DSS requires decision-makers to participate, man-machine dialogue is used to manipulate data through models. What is supported is only the structured

and clearly procedural part of the decision-making process. The core content of decision support system is human-computer interaction. In order to help decision-makers deal with semistructured and unstructured problems, identify objectives and environmental constraints, further clarify problems, generate decision schemes, and comprehensively evaluate decision schemes, the system should have stronger human-computer interaction ability and become an interactive system. As the decision-making environment becomes more complex, the limitations of DSS in decision support are becoming more prominent:

(1) Decision support system uses a static model to operate the data through the model. DSS is a system that makes comprehensive use of a large amount of data, organically combines many models, and assists decision-makers at all levels to realize scientific decision-making through human-computer interaction. The role of the system in decision support is passive and cannot be based on the decision environment. Changes provide active support [8, 9].

(2) Decision support system is modeled by the decision-makers and requires decision-making problems with procedural and clear computability [10, 11], which cannot support the unstructured problems that are common in decision-making.

(3) Decision support system is not a general product, but a solution. Each enterprise should combine its own situation, clarify the management difficulties to be solved, and then analyze, design, develop, and implement the decision support system, so as to truly meet the needs of enterprise management decision-making. DSS is based on quantitative mathematical models and lacks corresponding support methods for qualitative, fuzzy, and uncertain problems in decision-making [12, 13].

Intelligent decision support system (IDSS) is a combination of artificial intelligence (AI) and DSS and applies expert system (ES) technology to enable DSS to more fully apply human knowledge, such as descriptive knowledge about decision-making problems and procedural knowledge in the decision-making process. Introducing AI technology into DSS is mainly through the combination of expert system and DSS and adding inference engine and rule base to DSS system. In the decision-making process, many knowledge amounts cannot be expressed by data or described by models, so there is no fixed way of expertise and historical experience. Researchers integrate AI technology into DSS, mainly through the combination of expert system and DSS, and add inference engine and rule base to DSS system. The rule base introduced by IDSS can store these knowledge amounts and provide important reference and basis for decision-making. IDSS can have many types of information bases: text base (TB), database (DB), arithmetic base (AB), model base (MB), and rule base (RB). The text library stores a large number of documents written in natural language. The database stores the field form of

the key factors of things. Various models reflecting the essential relationship of information are stored in the model base. Rule base is the most refined form of knowledge. From the original unprocessed data to the processed information and then to the extracted knowledge, this evolution relationship of information is called evolution chain.

From the viewpoint of system level, IDSS can be technically divided into three levels:

(1) The application layer is directly oriented to IDSS users. In this layer, decision-makers can determine the status and constraints of IDSS according to their own needs. Decision-makers conduct system dialogue and input relevant information through the user interface; DSS understands user requests and commands through information conversion, carries out system reasoning operation, and reflects the results to users through the output interface. The whole process is transparent to users.

(2) Control and coordination layer, for the chief designer of IDSS: its basic unit is the control and coordination module of the system central library. The system engineer establishes the relationship between them through the standard interface of each library.

(3) The basic structure layer is for professional programmers. Professional programmers implement each library through this layer, specifically defining the organizational structure and communication mode of each library, so as to complete the department management and external communication tasks of each library.

During the operation of IDSS, each module needs to call the upper bridge repeatedly, which is less efficient than using the low-level call directly. However, considering that IDSS only operates when senior managers make major decisions, its operation frequency is much lower than that of other information systems, and the environmental conditions of each operation are very different; it is completely worth sacrificing part of the operation efficiency in exchange for the efficiency of system maintenance [14–16].

With the development of computer and artificial intelligence technology, the research focus of IDSS has gradually shifted from expert IDSS to the research of IDSS model system, man-machine interface, knowledge processing unit, and distributed IDSS. Distributed IDSS introduces distributed artificial intelligence (DAI) technology on the basis of knowledge-based IDSS. Its main idea is to use agent and multiagent system technology in Da; that is, agents playing different roles are designed as intelligent agents of modules according to the main functional modules in IDSS [17]. Through the self-learning mechanism, it simulates the different steps of human intelligence to complete the decision-making task, so as to make scientific decision-making. Distributed IDSS mainly designs and establishes large-scale and complex intelligent decision support system supported by Internet [18].

Expert system tools and technologies can be integrated into decision support system to provide users with consulting environment to improve decision quality and complete functions that conventional decision support system cannot complete. However, expert system is different from decision support system. When using decision support system, users must have certain professional knowledge and skills for the problems they deal with; that is, users should know how to reason about the problems, what questions should be put forward, how to get the answer, and how to proceed to the next step. At this time, the decision support system can only assist users to make decisions. However, the expert system is different. It itself has the professional knowledge of experts in a certain field to solve problems. Users only need to put forward the facts and appearances of the problem to the expert system.

3. System Model Design

3.1. Intelligent Decision Support System. Since the emergence of the intelligent decision support system, due to the great potential of expert system technology in the field of management decision-making, researchers at home and abroad have conducted a lot of research, integrating tools and methods such as computer science and artificial intelligence with human decision-making processes [19]. Intelligent decision system supports structure and function [20]. Holsapple summarized the system's support ability and learning ability for decision-making process and divided IDSS into 4 categories [21]: no adaptive, passive support; no adaptive, which can provide active support; adaptive, passive support; adaptive, active support.

3.1.1. Active Decision Support System. The traditional decision support system provides the corresponding data and model, and the user chooses the corresponding method and model. The decision process is completely controlled by the user. The system only completes the auxiliary computing function. Like many developing things, active decision support system (ADSS) is a developing concept. The intelligent decision support system has part of the domain expert knowledge. Human intelligence has the ability of actively supporting decision-making. ADSS can change its behavior with the user's decision-making process, which is an important milestone in the research of decision support system, which is generated by intelligent decision support system [7]. Active decision-making is an important feature of ADSS. By establishing a human cognitive model, ADSS can provide decision-makers with different choices in different problem solving stages, thus forming different problem solving paths. ADSS provides decision-makers with different method choices at different stages of decision-making problem solving by establishing human cognitive model. ADSS is based on human prior knowledge, but its premise is that the system runs in a static decision-making environment. Therefore, ADSS still has the limitation of poor adaptability in practical application.

However, the research on ADSS lays a foundation for the proposal of adaptive decision support.

3.1.2. Adaptive Decision Support System. ADSS relies mainly on human prior knowledge. The operating environment of the system is static. The domain knowledge and reasoning knowledge required for decision-making are known in advance. In fact, the decision-making environment is usually changeable, and the problem solving process is closely related to the decision-making environment. In relation to this, the knowledge that humans possess and can use for reasoning is also limited. Therefore, the above ADSS knowledge structure and function have many limitations. Adaptive decision support system is an important step toward better support decision. Adaptability is based on the environmental changes of the system and improves the ability of the system of dealing with problems. In order to overcome the limitations of ADSS in knowledge structure, data mining, data warehousing, case-based reasoning, and other data-driven decision support methods and machine learning techniques are adopted. Algorithms such as ANN, GA, and ROUGH SET attempt to discover knowledge that is critical to decision-making from a large amount of historical data and past experience, so that the system has the ability of adjusting its behavior over time and decision process changes, thus generating adaptive decision-making support system (ADSS) [22].

The researchers try to find the knowledge related to decision-making problems from a large number of historical data and past experience by using machine learning and case-based reasoning, so as to make the system have the ability of adjusting its behavior with the change of time and decision-making process. On this basis, people have carried out a lot of research on ADSS, including system structure adaptation, domain knowledge adaptation, and user interface adaptation. Adaptability and self-learning ability have become a main symbol of intelligent decision support system. ADSS has logic-based reasoning in addition to traditional process calculations and other forms of reasoning and also uses inductive reasoning to implement dynamic knowledge systems. Inductive reasoning is a kind of nonmonotonic reasoning, which can be limited or incomplete knowledge. The state introduces a complete state of knowledge. Through inductive learning, ADSS has a certain ability of innovating and can use inductive assertions as knowledge. When new contradictions are contradictory, the knowledge obtained by inductive reasoning can be overthrown, thus maintaining the consistency of knowledge [23].

3.1.3. Decision Expert System. The decision expert system is a decision support system established by using expert technology. The decision expert system uses deductive reasoning and uses existing knowledge to derive conclusions. The correctness of the derivation process can be guaranteed. The problem of this system is that a complete axiom system is needed for the basis of reasoning. In fact, in the environment of uncertain, abrupt, and fuzzy information, this condition is difficult to achieve, so it is only suitable

for the application of well-defined decision-making tasks. In recent years, the progress of artificial intelligence and expert system technology seems to break through the limitations of traditional decision-making expert systems; the application of nonmonotonic reasoning and qualitative reasoning technology is broadening the application scope of expert systems and has made some progress in combination with human intelligence [24].

3.1.4. Holistic Decision Support System. The integrated decision support system is based on the adaptive decision support system and the decision expert system [7]. In the process of human expert decision-making, it faces incomplete, uncertain, and even conflicting knowledge, and human thinking often has nonprocesses. Sexuality generally leads to decision-making through the synthesis of various knowledge and processes. The integrated decision support system (HDSS) emerges to mimic human advanced intelligence and can make full use of human in process analysis, logical reasoning, and cognition and learning. And it has the advantages of knowledge innovation, so that the system's auxiliary decision-making ability transcends the stage of fact, reasoning, and learning and can support the decision-making problem of ill-structured structure [25].

The application of Internet technology in the field of decision support makes the decision environment have new characteristics; that is, the data in decision analysis are no longer concentrated in one physical location, but scattered in different departments or regions. The above discussion about the types of decision support systems is based on the main characteristics of the system, from the perspective of the development of system intelligence. Several types are mutually contained and complementary. More advanced models reflect the evolution of system intelligence [26, 27], Excluding other types of features. In the big data environment, distributed decision support system will get more and more attention. In fact, the research of intelligent decision support systems is showing a trend of integration [7]. Existing systems are generally hybrid systems formed by a combination of methods, with several models characteristics. Uncertainty is the key problem in the current research of artificial intelligence technology, and it is also the core problem throughout the whole process of big data intelligent decision-making. Table 1 compares several IDSSs from the aspects of system learning ability, intelligent behavior, and decision-making methods.

3.2. Intelligent Information System Requirements Analysis. Since the beginning of the 21st century, the wave of globalization has driven the rapid development of the social economy, and the quality of human life has been continuously improved. Together with the rapid development of information technology, human beings have begun to pursue the information technology and intelligence of the living environment. Such powerful human subjective needs have spawned the birth of intelligent information systems. Although the intelligent information products have been developed over a long period of time, this paper has found

TABLE 1: Comparison of IDSS models.

System	Knowledge base type	Completeness of knowledge	Knowledge innovation ability	Decision process understanding	Main reasoning	Main decision-making tool
Active decision support system	Static	Completeness without conflict	No	Have	Adopt a predefined process	Data, cognitive model
Decision expert system static	Static	Completeness without conflict	No	Maybe do not have	Deductive reasoning	Knowledge base
Adaptive decision support system	Dynamic	Allowing incompleteness and conflict	Yes	Generally have	Deductive and inductive reasoning	Data, intelligent decision model, and knowledge base
Integrated decision support system	Dynamic	Allowing incompleteness and conflict	Yes	Have	Deductive, inductive, and case-based reasoning	Cognitive model, intelligent computing method, machine learning, and knowledge base

that the existing smart products or systems still have the following shortcomings after investigating the existing products or systems in the market:

(1) Manufacturer's products are singularly mass-produced.

The existing smart device manufacturers only carry out mass production of a single product and do not systematically carry out product design and mass production and do not form a unified industry standard or access standard resulting in a mixed market of products and systems. There is poor compatibility between the two; new products are difficult to access the system.

(2) The system control layer does not achieve true intelligence.

In response to the above problems (1), many manufacturers have proposed their own intelligent system solutions, integrating different types of products into the system and providing decision control functions, realizing remote control, security alarm, and other functions. However, the decision-making control layer of the system is not intelligent, but the user behavior is used for control or decision-making; that is, the user needs to manually adjust it.

(3) The pressure on data processing of system servers in the Internet of Things has increased dramatically.

Most of the existing intelligent products use the intelligent gateway for data forwarding, and all the original data is sent to the remote server for analysis, storage, processing, etc. After the user scale is expanded, all the user's original data is processed by the remote server. The pressure on servers is getting bigger and bigger, and the cost of expansion is getting higher and higher.

In view of the shortcomings of the above existing intelligent products or systems, the intelligent system studied in this paper hopes to make improvements and finally build a gateway that can provide a unified device interface, realize decision control intelligence, and include local data storage and processing functions. The intelligent information system of the device provides users with systematic, automated, intelligent, and personalized services.

3.3. Hierarchical Structure Design of Intelligent Information System. The structure of the Internet of Things itself is complex and diverse. The current structure of an Internet of Things is divided into three levels: the perception layer, the network layer, and the application layer.

The bottom layer of the Internet of Things is the perception layer, which is the basis for realizing the comprehensive perception of the Internet of Things. RFID, sensors, two-dimensional code, etc., are mainly used to collect device information with sensors and use radio frequency identification technology to achieve transmission and recognition within a certain range. The main function is to identify objects and collect information through sensing devices.

The network layer is located above the sensing layer and is a network device and platform that serves the aggregation, transmission, and preliminary processing of IoT information. Through the existing three-network or next-generation network NGN, the huge amount of data collected from the sensor network is seamlessly transmitted over long distances; it is responsible for the secure transmission of the information collected by the sensor and the analysis of the collected information, processing and providing the results to the application layer.

The application layer is the top layer of the Internet of Things architecture. It mainly solves the problem of information processing and human-machine interface and provides the information services that people need through data processing and solutions. The application layer directly contacts the user and provides users with rich service functions. The user customizes the required service information, such as query information, monitoring information, and control information, on the application layer through the smart terminal.

The application of Internet of Things technology in the service environment mainly focuses on home appliance

automation and intelligent security systems. Home appliance automation is installing sensors in traditional household electrical appliances (such as air conditioners, TVs, and refrigerators), making them intelligent, nodding them, and accessing the Internet, so that users can realize remote control of home appliances in the sky. The intelligent security system is equipped with sensor nodes for monitoring fire, gas concentration, etc., at home, so that the user gets the first time alarm feedback when there is a fire in the home and can timely handle the danger or escape in time to avoid personnel casualties.

The basic framework of an intelligent information system can be divided into three layers: the environment-awareness layer, the network transmission layer, and the smart service application layer as shown in Figure 1.

3.3.1. Environment-Aware Layer.
The environment-aware layer is the bottom layer of the system. If an analogy is made by one person, the environment-aware layer is like a human tactile nerve, which is composed of a large number of environmental information sensors. These sensors are like a neural node, distributed in every corner of the environment, collecting environmental information and sending the collected data to the upper layer of the intelligent information system through a certain short-range wireless transmission technology.

From the perspective of network technology, the sensor of the environment-aware layer constructs a WPAN network within a certain range. The so-called WPAN is a network proposed to solve the "last few meters of wireless communication connection". Generally, this refers to short-range wireless networks with coverage within a radius of 10 m, especially for self-organizing networks that can be connected shortly between portable consumer appliances and communication devices. WPAN is a network that is parallel with wireless wide area network (WWAN), wireless metropolitan area network (WMAN), and wireless local area network (WLAN) but has a smaller coverage. The corresponding relationship is shown in Figure 2.

3.3.2. Network Transport Layer.
The network transport layer is the middle layer of the intelligent information system of the service environment, which assists network communication and data transmission between the entire system levels. It includes communication between environmental information collection sensors, communication between environmental information collection sensors and terminals, and communication between terminals and smart cloud platforms.

3.3.3. Smart Service Application Layer.
The intelligent service application layer is the uppermost layer of the system and is the concentrated expression of the intelligence of the information system oriented to the service environment. It mainly includes the following functions:

(a) Cloud data storage and management

(b) Inducting user habits and building user knowledge base

(c) Smart decision based on user habits

(d) Sending control commands to the lower layer of the system

If one person is used as an analogy, the intelligent service application layer is equivalent to the human brain, which can store memory, intelligently think, make optimal decisions, and control other parts of the human body through neural networks. The intelligent service layer learns knowledge through machine learning algorithms and makes decisions to control the underlying devices of the system.

3.4. Improvement of Data Fusion Algorithm.
The limitations of the fuzzy fusion algorithm in the current environment are mainly reflected in the two aspects of time efficiency and energy consumption. The improvement of the algorithm will also start from these two aspects: first, it reduces the amount of data to be fused by setting a reasonable threshold to improve the fusion algorithm by setting weights for data that fails to pass the threshold.

In an intelligent information system, assuming that there are n sensor nodes, they can form $2n - 1$ sensor groups. As the number of nodes n increases, the number of sensor groups increases exponentially. Each sensor group performs the calculation of the tolerance function and the fuzzy measure function, which will generate a huge amount of computation and take a lot of time. Correspondingly, if we can reduce the number of n participating in the operation by a certain method, the calculation amount and time consumption of the algorithm will decrease exponentially. Therefore, reducing the amount of data with fusion is an effective way to solve the above problems.

Assuming that the system enters a stable data transmission phase, if the monitoring index does not fluctuate significantly over a period of time, then this part of the data is relatively small for system decision-making. The method designed in this paper is to record the result of each fusion and set an appropriate threshold. The data to be transmitted is compared and decided to be retained or deleted. The results of the reservation are directly involved in the final fusion, and the data to be deleted is given a weight, giving it the opportunity to reengage. The main process of the algorithm is as follows:

Suppose that the output set of n sensors is

$$X = [x_1, x_2, \ldots, x_n]. \tag{1}$$

After performing a fusion, the result of the fusion is recorded as C_0, and, in addition, the threshold is set as

$$Thr = [C_0 - \mu, C_0 + \mu]. \tag{2}$$

In the next fusion, the node output x_i is compared with Thr. If $x_i \in Thr$, the data is deleted and the change output is $x_1, x_2, \ldots, x_m, m \leq n$. There are three key points here: the selection of thresholds, the processing of deleted data, and the improvement of weight-based algorithms.
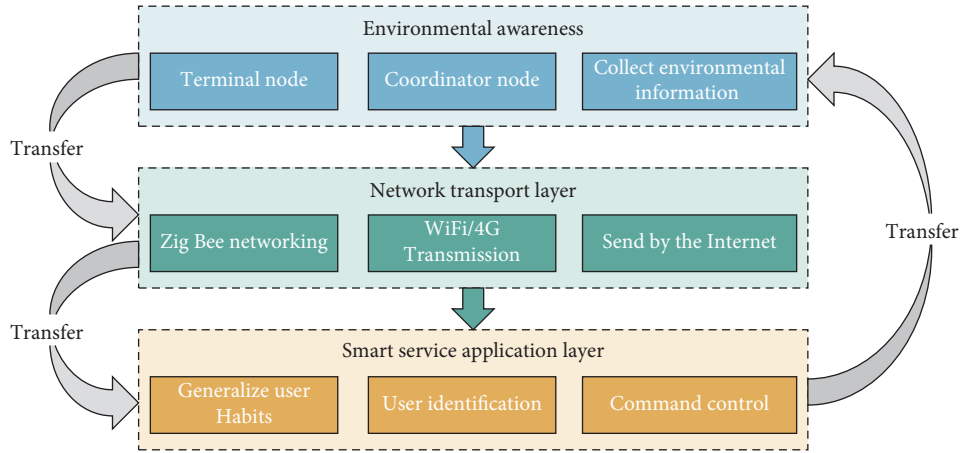
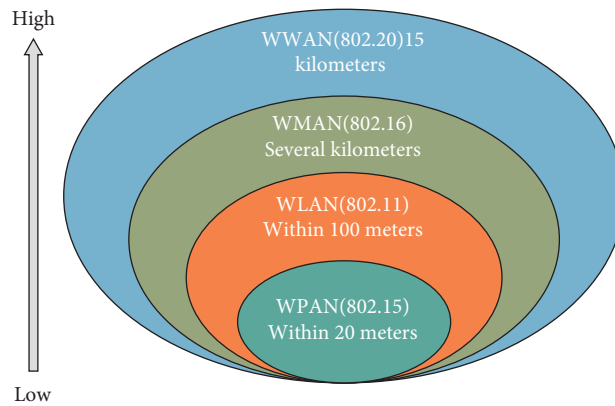FIGURE 1: Intelligent information system hierarchy diagram.

FIGURE 2: Correspondence between WPAN and other wireless networks.

3.4.1. Threshold Selection. Choosing the appropriate threshold is especially important in the improvement of the algorithm. If the threshold is selected to be too large, it will not be able to mask invalid data, and the accuracy of the data will be greatly reduced. If the threshold is selected to be too small, it will cause the data to be deleted and screened out in the next comparison, which will result in the algorithm. The amount of calculation and time consumption are not much different from those before the threshold is taken, and the meaning of the improved algorithm is lost. According to the above formula (2), the selection of the threshold in this paper is based on the result of the previous fusion C0. Here, we consider the fusion of temperature data.

Assume the following scenario: after the user enters the service environment, the room temperature is 8°C. The intelligent information system sets the air conditioner temperature to 25°C according to the user's habit. After the air conditioner is turned on, the room temperature rise curve is shown in Figure 3.

From the figure, we can see that, in the first few minutes, since the main engine has just started running after the air conditioner is started, the heating effect is not obvious and the temperature is almost changed. With the normal operation of the air conditioner main unit, the heating effect

begins to appear and the temperature rises faster; later, as the room temperature has risen, the temperature rise requires more heat, so the temperature rise tends to be slow.

In this scenario, the fuzzy fusion of the temperature data is carried out. Obviously, the data in the first few minutes and the last few minutes shows a lot of redundancy, and some data can be eliminated by setting the threshold. After repeated experiments and measurements, this paper selects $\mu = 0.5$, which is the threshold of

$$Thr = [C_0 - 0.5, C_0 + 0.5]. \tag{3}$$

3.4.2. Delete Data Processing. In the above threshold elimination process, the data within the threshold range is directly deleted, but if the amount of data deleted is large, it will also cause distortion of the final fuzzy fusion result. In order to prevent this from happening, this paper introduces a processing mechanism for deleting data, which is to give the data a weight value to reflect its importance in the system.

The weight value mechanism of this paper is to assign a weight W to each data to be fused. The default value is
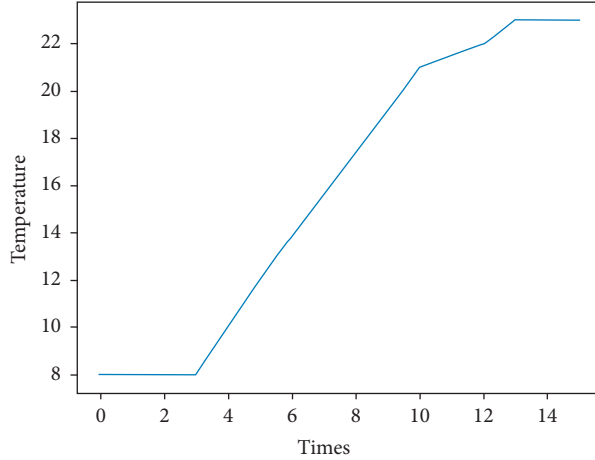
Figure 3: The room temperature rise curve.

$W = 1$. W represents the number of data amounts similar to the data and $W = 1$ means that the data can only represent itself.

According to the above, assuming that the result of the last fuzzy fusion is C_0, a weighted data to be merged x_i is compared with the threshold value (3), and if $x_i \in Thr$, whether it is the first data entering the threshold Thr, then, C_0 is assigned to x_i, and the weight W is still 1; the data that enters the threshold Thr again can be directly deleted, but each time a data enters the threshold the weight W of C_0 is increased by 1, indicating that the number of data amounts similar to C_0 is increased by one. If $x_i \notin Thr$, the data is not adjusted, participating in the next fusion, and the weight W is still 1. This process can be represented in Figure 4.

3.4.3. Weight-Based Algorithm Improvement. Since the weight of the data to be merged is introduced, the fuzzy measure function of the traditional fuzzy fusion algorithm cannot be directly used, and corresponding improvement is needed. The following describes the improvement of the fuzzy measure.

The fuzzy measure function reflects the reliability of the sensor group participating in the calculation. The range is between [0, 1]. The more the number of sensors participating in the fusion, the larger the fuzzy measure value and the higher the reliability. Otherwise, the number of participating sensors is less than the fuzzy measure value and thus the reliability is lower.

Assume, that in the home service environment, the number of sensors in the sensor group is n. When the number of participating sensors is m when a data is fused, the reliability can be expressed as m/n.

The result of the previous fusion is C_0. After the next fusion, all data is compared with the threshold T, and the weight of C_0 is $W = i$.

If C_0 is not included in the sensor group data to be calculated, m/n is still used as the result of the fuzzy measure. If C_0 is included in the sensor group data to be calculated, the fuzzy measure of the sensor group can be expressed by

$$\frac{(m + i - 1)}{n}. \tag{4}$$

The improved fuzzy measure guarantees the integrity of the data and makes the fusion result more accurate.

4. Designs and Analysis

4.1. System Architecture Design. This paper participates in the design of the presidential architecture of a service-oriented intelligent information system, as shown in Figure 5.

The intelligent information system is divided into functional modules, and the specific functions are designed. The functional modules include information collection template, data processing module, knowledge management module, decision control module, and device adjustment module. Their relationship to each other is shown in Figure 6.

The data collected by the information acquisition module is the most primitive data prototype obtained by the data processing module. It is the nerve ending of the entire intelligent information system and collects the information of the surrounding environment as accurately as possible for system processing, calculation, and decision-making.

The data processing module is a template for the intelligent information system. First, data is collected as input, then stored locally, and then preprocessed to provide a more accurate and reliable data prototype for the knowledge management module and decision control module.

The knowledge management module is a positioning in the intelligent information system, which is equivalent to a memory module in the human brain. Its input mainly comes from the data prototype provided by the data processing module, and the output is the user habit obtained by massive data accumulation and data mining. These data are stored in the database together with the data prototype and, on the other hand, as decision control. The data also can be input to the module for reference by the decision control module.

The decision control module is mainly responsible for making comprehensive decision judgments based on the information transmitted by the data processing module and
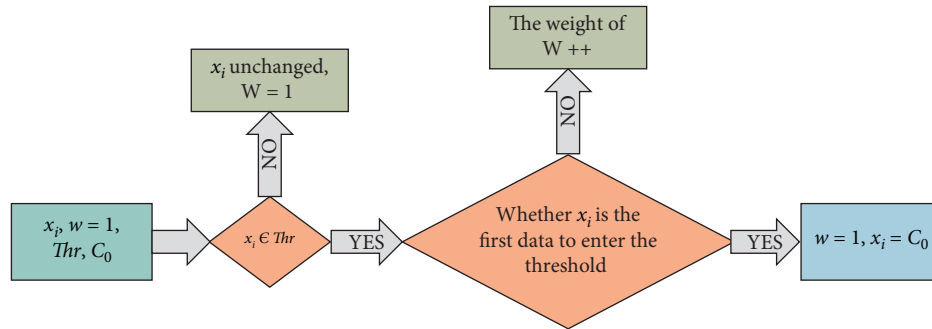
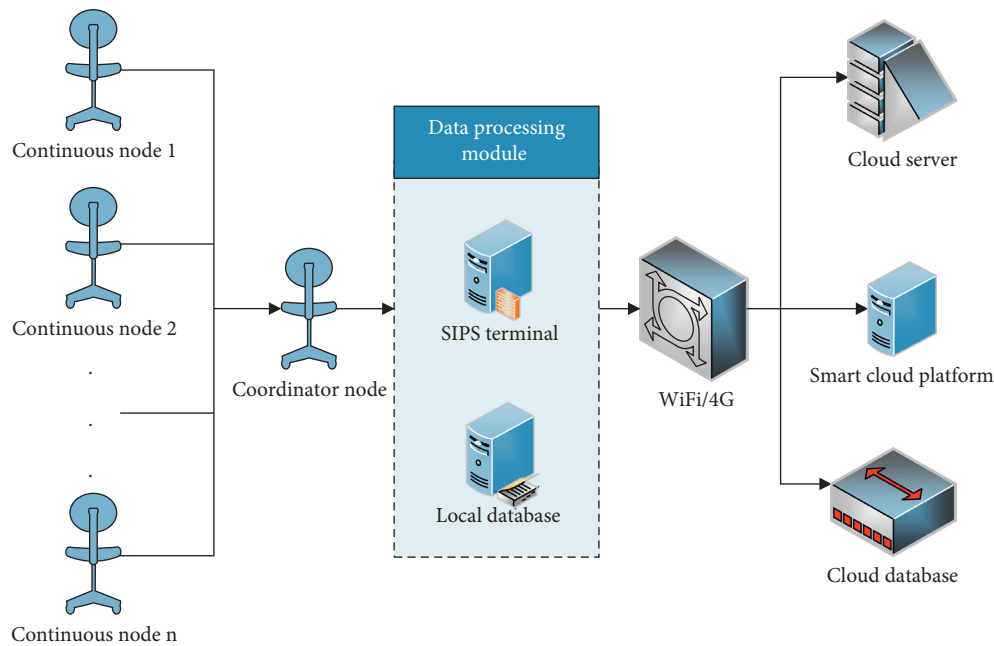FIGURE 4: Data processing process that introduces weight values.

FIGURE 5: Intelligent information system architecture diagram.

the knowledge management module and producing control instructions.

The device adjustment template is one of the ends of the intelligent information system, receives the control command of the decision control module as an input, and performs input information to control the device parameters.

4.2. Experimental Analysis of Improved Data Fusion Algorithm. In this section, by setting the specific application scenarios, the data is simulated by MATLAB, and the traditional fuzzy fusion algorithm and the improved algorithm are compared and analyzed.

4.2.1. Comparison of the Number of Participating Nodes. First of all, this paper separately compares the number of sensors participating in each fusion of the two algorithms and obtains the results of Figure 7 as follows.

In Figure 7, the blue dot curve represents the traditional data fusion algorithm, and the orange dot curve represents the improved fuzzy fusion algorithm. As can be seen from the figure, in the traditional algorithm, the number of nodes participating in the fusion has been fluctuating around 90~100; this means that, in every fusion process, almost all the data collected by the nodes will participate, which will cause huge calculations and consume system resources and time. The improved fusion algorithm has dropped significantly by half since the second fusion and then basically maintains a stable fluctuation around 50. This shows that, from the second time, the number of nodes participating in the fusion is significantly reduced. That is, setting a reasonable threshold makes the number of nodes participating in the fusion greatly reduced. From the previous analysis, we already know that when there are n nodes participating in the fusion, $2n - 1$ kinds of node combinations will be generated. When n is reduced, the number of node combinations will decrease exponentially, which will greatly reduce the calculation amount of the system and improve
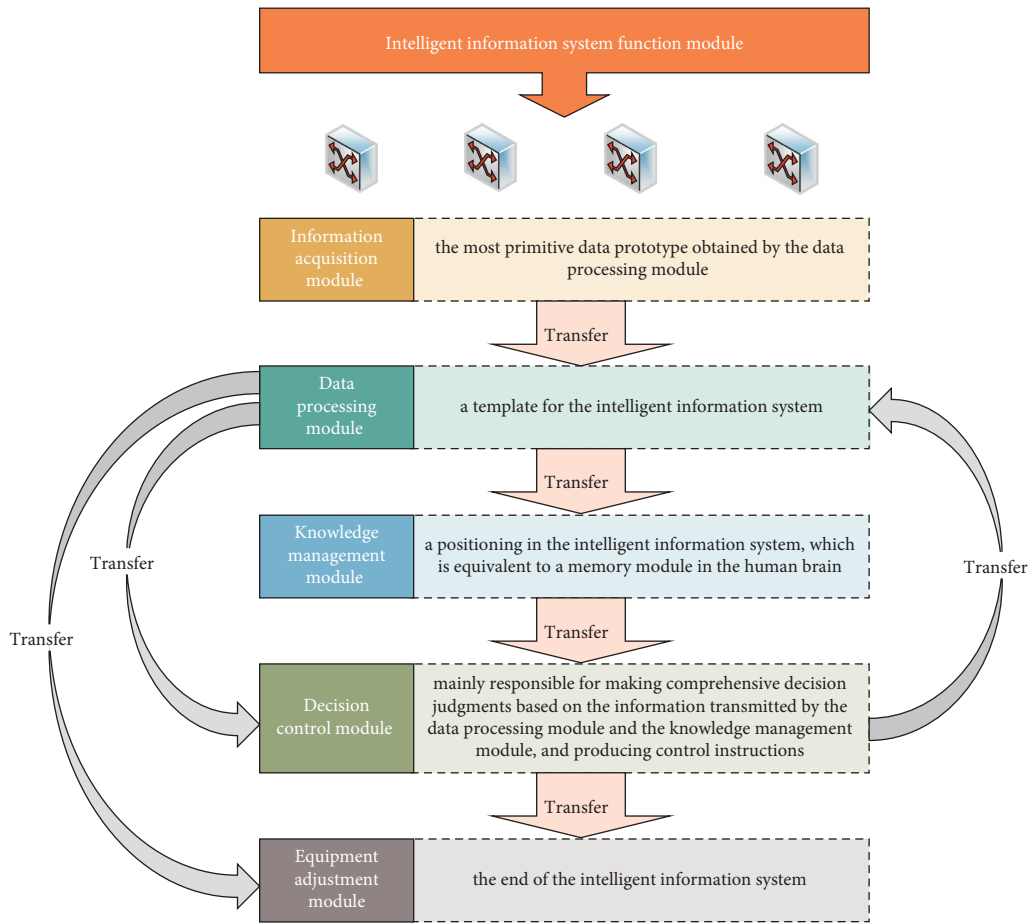
FIGURE 6: Intelligent information system function module diagram.
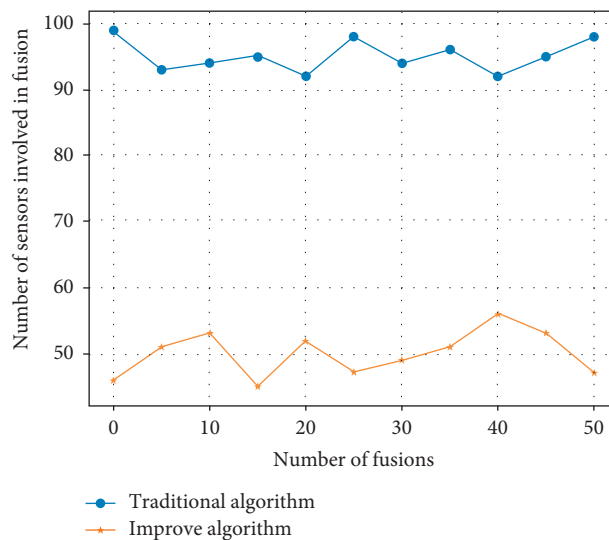
FIGURE 7: Comparison of the number of participating nodes.

the timeliness of data processing. Therefore, in the improved fuzzy fusion algorithm, it is feasible and effective to reduce the amount of data to be fused by setting a reasonable threshold.

4.2.2. Comparison of Fusion Results. At the same time, the fusion results of the traditional fuzzy fusion algorithm and the improved weight-based fuzzy fusion algorithm are compared. The comparison results are shown in Figure 8.
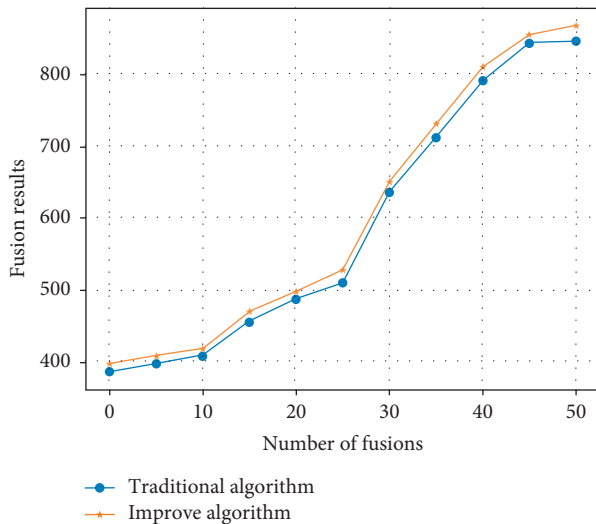
FIGURE 8: Comparison of fusion results.

In Figure 8, the blue dot curve represents the temperature fusion result of 50 fusions by the traditional fuzzy fusion algorithm, and the orange dot curve represents the temperature fusion result of the improved weight-based fuzzy fusion algorithm. It can be seen from the figure that the curve of the improved algorithm is more consistent with the curve of the improved algorithm and the error is smaller. This shows that the improved algorithm can still accurately integrate the data and the improved algorithm is accurate and reliable.

5. Conclusions

Since the introduction of the concept of intelligent decision-making information system, after years of research and development, it has gradually entered the practical stage of the market. With the rapid development of technologies such as Internet of Things, big data, and cloud computing, its research has gradually focused on the wisdom of the system level. In view of the shortcomings of existing intelligent products or systems, this paper designs a data processing module of intelligent information system, considers the problem of data fusion after system expansion, and improves the data fusion algorithm. According to the analysis of requirements, the intelligent information system is designed, and the three-layer architecture of the system based on intelligent decision-making is proposed, including the sensing layer, the network transmission layer, and the intelligent service application layer. The module construction is decomposed to explain how the modules interact with each other to cooperate in work. Considering the massive data problem after the expansion of intelligent information system, the data fusion algorithm is compared and improved according to the traditional fuzzy fusion algorithm. A weight-based fuzzy fusion algorithm is proposed to make it more adaptable to massive data and improve the accuracy and reliability of system data. Through the simulation diagram, the improved algorithm and the traditional algorithm are compared and analyzed, and the effectiveness of the weight-based fuzzy fusion algorithm is verified.

References

[1] V. Vemulapalli, J. Qu, J. M. Garren et al., "Non-obvious correlations to disease management unraveled by bayesian artificial intelligence analyses of CMS data," *Artificial Intelligence in Medicine*, vol. 74, pp. 1–8, 2016.

[2] H. Xiong, X. Wang, and Y. Zhu, "Modeling and verification of real-time interaction for IEC 61850 intelligent electronic device," *Automation of Electric Power Systems*, vol. 38, no. 19, pp. 90–95, 2014.

[3] M. Dzikovska, N. Steinhauser, E. Farrow, J. Moore, and G. Campbell, "Beetle II: deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics," *International Journal of Artificial Intelligence in Education*, vol. 24, no. 3, pp. 284–332, 2014.

[4] H. Zhu, "Coordination innovation architechture for IoT and development strategy of smart service industry," *Journal of Nanjing University of Posts & Telecommunications*, vol. 34, no. 1, pp. 1–9, 2014.

[5] M. Brambilla, E. Umuhoza, and R. Acerbis, "Model-driven development of user interfaces for IoT systems via domain-specific components and patterns," *Journal of Internet Services and Applications*, vol. 8, no. 1, p. 14, 2017.

[6] M. C. Bell, "Environmental factors in future transport," *Cuestiones Constitucionales*, vol. 59, no. 18, pp. 349–356, 2010.

[7] X. Sun, Q. Li, and C. Sun, "Research on fuzzy adaptive intelligent decision-making in complex environment," *Journal of Physics: Conference Series*, vol. 1345, no. 5, Article ID 052085, 2019.

[8] G. Phillips-Wren, M. Mora, G. A. Forgionne, and J. N. D. Gupta, "An integrative evaluation framework for intelligent decision support systems," *European Journal of Operational Research*, vol. 195, no. 3, pp. 642–652, 2009.

[9] P. Rydahl, L. Hagelskjaer, and L. Pedersen, "User interfaces and system architecture of a web-based decision support system for integrated pest management in cereals," *Eppo Bulletin*, vol. 33, no. 3, pp. 473–481, 2010.

[10] S. C. Fan and K. C. Yu, "How an integrative STEM curriculum can benefit students in engineering design practices," *International Journal of Technology & Design Education*, vol. 27, no. 1, pp. 1–23, 2015.

[11] D. Longinotti, "Computationalism and the locality Principle," *Minds and Machines*, vol. 19, no. 4, pp. 495–506, 2009.

[12] R. Caruana, M. J. Carrington, and A. Chatzidakis, "Beyond the attitude-behaviour gap: novel perspectives in consumer

ethics": introduction to the thematic symposium," *Journal of Business Ethics*, vol. 136, no. 2, pp. 215–218, 2016.

[13] J. F. Baldwin, "Support logic programming," *International Journal of Intelligent Systems*, vol. 1, no. 2, pp. 73–104, 2010.

[14] P. F. Butzen, V. Dal Bem, and A. I. Reis, "BTI and HCI first-order aging estimation for early use in standard cell technology mapping," *Microelectronics Reliability*, vol. 53, no. 9-11, pp. 1360–1364, 2013.

[15] A. R. Simmons and K. B. R. Baggerly, "The emerging role of HE4 in the evaluation of epithelial ovarian and endometrial carcinomas," *Oncology*, vol. 27, no. 6, pp. 548–556, 2013.

[16] R. Seidl, C. S. Eastaugh, and K. Kramer, "Scaling issues in forest ecosystem management and how to address them with models," *European Journal of Forest Research*, vol. 132, no. 5-6, pp. 653–666, 2013.

[17] J. E. Hunink, I. A. Niadas, P. Antonaropoulos, P. Droogers, and J. De Vente, "Targeting of intervention areas to reduce reservoir sedimentation in the Tana catchment (Kenya) using SWAT," *Hydrological Sciences Journal*, vol. 58, no. 3, pp. 600–614, 2013.

[18] Q. Zhou, "Research on heterogeneous data integration model of group enterprise based on cluster computing," *Cluster Computing*, vol. 19, no. 3, pp. 1275–1282, 2016.

[19] Q. Zhou and J. Luo, "The study on evaluation method of urban network security in the big data era," *Intelligent Automation and Soft Computing*, vol. 24, no. 1, pp. 133–138, 2018.

[20] Q. Zhou, "Multi-layer affective computing model based on emotional psychology," *Electronic Commerce Research*, vol. 18, no. 1, pp. 109–124, 2018.

[21] T. Li and R. J. Kauffman, "Adaptive learning in service operations," *Decision Support Systems*, vol. 53, no. 2, pp. 306–319, 2012.

[22] N. F. Matsatsinis, "CCAS: an intelligent decision support system for credit card assessment," *Journal of Multi-Criteria Decision Analysis*, vol. 11, no. 4-5, pp. 213–235, 2002.

[23] M. Gahegan, "On the application of inductive machine learning tools to geographical analysis," *Geographical Analysis*, vol. 32, no. 2, pp. 113–139, 2000.

[24] W. Ouerdane, N. Maudet, and A. Tsoukiàs, "Argumentation theory and decision aiding," *International Series in Operations Research & Management Science*, vol. 142, pp. 177–208, 2010.

[25] S. Jiménez, T. De La Rosa, and F. Susana, "A review of machine learning for automated planning," *The Knowledge Engineering Review*, vol. 27, no. 4, p. 35, 2012.

[26] Q. Zhou, Z. Xu, and N. Y. Yen, "User sentiment analysis based on social network information and its application in consumer reconstruction intention," *Computers in Human Behavior*, vol. 100, pp. 177–183, 2019.

[27] Q. Zhou, Z. Zhang, and Y. Wang, "Research on safety management system optimization of B2C e-commerce intelligent logistics information system based on data cube," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 2, pp. 1585–1592, 2020.

Research Article

Predictive Control-Based Completeness Analysis and Global Calibration of Robot Vision Features

Jingjing Lou (iD)

School of Mechanical and Electrical Information, Yiwu Industrial and Commercial College, Yiwu, Zhejiang 322000, China

Correspondence should be addressed to Jingjing Lou; chinglou@ywicc.edu.cn

This paper provides an in-depth study and analysis of robot vision features for predictive control and a global calibration of their feature completeness. The acquisition and use of the complete macrofeature set are studied in the context of a robot task by defining the complete macrofeature set at the level of the overall purpose and constraints of the robot vision servo task. The visual feature set that can fully characterize the macropurpose and constraints of a vision servo task is defined as the complete macrofeature set. Due to the complexity of the task, a part of the features of the complete macrofeature set is obtained directly from the image, and another part of the features is obtained from the image by inference. The task is guaranteed to be completely based on a robust calibration-free visual serving strategy based on interference observer that is proposed to complete the visual serving task with high performance. To address the problems of singular values, local minima, and insufficient robustness in the traditional scale-free vision servo algorithm, a new scale-free vision servo method is proposed to construct a dual closed-loop vision servo structure based on interference observer, which ensures the closed-loop stability of the system through the Q-filter-based interference observer, while estimating and eliminating the interference consisting of hand-eye mapping model uncertainty and controlled robot input interference. The equivalent interference consisting of hand-eye mapping model uncertainty, controlled robot input interference, and detection noise is estimated and eliminated to obtain an inner-loop structure that presents a nominal model externally, and then an outer-loop controller is designed according to the nominal model to achieve the best performance of the system dynamic performance and robustness to optimally perform the vision servo task.

1. Introduction

The ability of humans to observe things in the outside world with both eyes and manipulate what we see is the result of a long evolutionary process that is important for us to adapt to and make use of the outside world. Inspired by this phenomenon, visual sensing was introduced to robot control systems to sense the state of the system and the task. This type of information is collected from the outside of the control system, compensating for the information fed back from within the system, relaxing the requirements for mechanical accuracy and joint tightness of the robot components, and allowing the robot system to enhance its flexibility in an uncertain external environment for various automatic control functions, such as robotic assembly, robotic surgery, and remote-control tasks. As a result, visual

servo control, that is, the introduction of visual information into the control closed loop to guide dynamic system motion, has received a great deal of attention. The development of robotics is changing rapidly; robots are more functional and more versatile; not only traditional robots for industrial applications but also more military and medical robots along with the wave of artificial intelligence robots are making a splash in various fields [1]. Whether the information contained in the macro- and microfeatures is complete affects the completion of the task. Therefore, the complete visual feature set of the visual servo task is defined to describe the feature set that meets the needs of completing the task. The feature set consists of a complete macrofeature set and a complete microfeature set. The wide application of robots makes robotics products form a huge industry. At the same time, the development of robotics will have a huge and far-

reaching impact on the country's comprehensive national power and its ability to develop sustainably. Therefore, robotics is not only seen as a frontier technology but also taken as a national strategy by many developed countries in the world, which gives strong support to the research in this field. As the most important class of robots, mobile robots will be used in industry, agriculture, national defines, scientific research, and other industries to partially or completely replace the work of people as human society continues to develop and the degree of information technology and intelligence continues to increase.

Vision-based mobile navigation is a research hotspot, and research in this area can currently be divided into three parts: visual simultaneous localization and mapping (VSLAM), path planning, and navigation control. Over the past decades, a large amount of research has focused on VSLAM, which is the study of environmental modelling. The process requires the system to build the environment while using the known local environment to achieve self-localization and then incrementally model the environment [2]. Research in this area, often without considering the command generation for its movement control, but only how to achieve environment reconstruction as well as real-time localization is given the information, is essentially a real-time version of structure from motion (SfM) in the field of computer vision [3]. In the visual navigation process, the captured image sequences are used to achieve online incremental environment perception and autonomous localization using VSLAM. The concept of feature completeness for robot vision servo tasks is proposed, and the complete feature set is defined to consist of a complete microscopic feature set and a complete macroscopic feature set. For an arbitrary vision servo task, a criterion is given to judge whether a microscopic feature set is complete or not, and three metrics are proposed to evaluate the performance of a complete microscopic feature set. The task of grasping a brush by a humanoid robot is taken as an example to study the acquisition and application of the complete microfeature set in the robot vision servo task [4]. According to the proposed definition and discriminative quasi-discrimination of the complete feature set, the complete feature set is designed. The camshaft algorithm is used to detect target points and actuator ends in the video stream captured by the system camera, and then the complete feature set is obtained.

Mobile robots often work in unstructured and unknown environments and need to autonomously sense and perceive the environment according to specified task goals and plan their behaviour efficiently and intelligently to accomplish specific tasks. To accomplish tasks more efficiently, mobile robot systems contain a central decision-making system and several subsystems that are responsible for different subtasks [5]. Among them, autonomous navigation technology for mobile robots is a fundamental module. The navigation system is the minimum complete system integrating environment modelling, scene understanding and reasoning, autonomous decision-making and execution, and human-robot interaction, corresponding to a complete processing process of information representation, association, fusion, reasoning, and management. Therefore, the level of

intelligence of a robot navigation system also determines the complexity of the environment it can adapt to and the complexity of the navigation task. The intelligence level of robot autonomous navigation systems is closely related to the corresponding methods of information processing, feature expression, and autonomous cognition. In recent years, the leap in technologies such as high-speed cameras, image processing hardware, and high-performance mobile computing platforms has provided mobile robots with the assurance of perceptual information and computing efficiency for autonomous navigation. The rapid development of information intelligence processing technology also keeps pushing the robot autonomous navigation system towards the direction of intelligence, which makes the robot gradually improve its adaptability and autonomy to the environment. This paper focuses on the effect of different feature sets on the completability and completion performance of an uncalibrated visual servo task. Feature completeness is proposed to examine whether a feature set provides enough information to make a task complete. Evaluation metrics are proposed to examine whether a completeness feature set enables optimal task completion. The humanoid robot grasping brush task and the robot writing Chinese calligraphy task are used as examples to investigate how to select, extract, and use the complete feature set for different visual servo tasks, respectively. Based on this, a robust calibration-free visual servo control method based on the complete feature set is proposed to improve the task completion performance.

2. Status of Research

Since the control commands of the vision servo system are calculated based on the hand-eye mapping relationship, the transient performance, steady-state performance, and stability of the system are affected by the hand-eye relationship [6]. The calibration-free vision servo system uses visual information to guide the robot's motion when the camera parameters are not calibrated. How to obtain an accurate hand-eye relationship in uncalibrated vision servo control is an important research problem [7]. The hand-eye mapping relationship for visual servo systems is strongly nonlinear and can be approximated by a series of linear relationships within each local domain, such as the use of image Jacobian matrices. There are two main ways to calculate the image Jacobian matrix, indirectly or directly, to estimate the matrix take values. One way to estimate the image Jacobian matrix indirectly is by estimating the unknown parameters in the model. The literature [8] represents the uncalibrated visual servo problem as a multiple-input, multiple-output adaptive control problem and accomplishes visual servo control using a Lyapunov adaptive control method based on SDU decomposition with online calibration of the camera parameters through a cascade control strategy. The literature [9] investigates how a system with unknown camera and robot parameters can accomplish calibrated control. The algorithm uses a depth-independent image Jacobian matrix to establish a mapping of visual deviations to the joint angle space of the robot arm so that the unknown camera and

robot parameters appear linearly in the closed-loop dynamic structure of the system, which in turn estimates the unknown parameters online using an adaptive algorithm [10]. Another approach estimates the image Jacobian matrix directly using a different method.

The spatial projection of image pixel points using multiview geometry as well as feature point matching and optical flow tracking is used to obtain a spatial point cloud, and the matching of this point cloud map with the current image and the PnP algorithm is used to localize the robot system [11]. Its localization optimization is performed with the reprojection error or image error as the optimization term for the pose solution. The method is characterized by high operable accuracy and intuitive map representation [12]. The basic dual-view geometric relative pose solution usually uses feature points to represent the image and calculates the basis matrix or single response matrix between two frames by matching the feature points and then obtains the relative pose between two frames by decomposing the singular values of the matrix [13]. The method has a deep theoretical foundation and can obtain pose calculation results with high accuracy under the premise of having ideal data matching [14]. This kind of reprojection method is usually implemented based on feature points, so this kind of method relies on the information of individual pixels in the image for localization, and this kind of matching method with local information often limits the reliability of data matching due to the limitation of its observation window size; in addition, the amount of information used in the selection and expression of pixels is limited, so this method has a weak ability to overcome changes in environmental conditions. In addition, due to the limited amount of information used, the ability to overcome changes in environmental conditions is weak, resulting in problems such as data matching errors or insufficient effective matching [15]. These problems also lead to the inability of visual navigation based on these methods to achieve reliable localization output and stable navigation under a long period or harsh working environment, which reduces the robustness of the visual navigation system.

Real-time modelling for the unstructured environment in the human-robot collaboration scenario can provide global process information for subsequent motion planning, which is a key technology for visual perception. It mainly studies the influence of different feature sets on the accomplish ability and completion performance of an uncalibrated visual serving task. The feature completeness is proposed to investigate whether a feature set provides enough information to make a task complete. According to the above comparison, the depth camera based on the principle can calculate the depth information in real-time and fast, with moderate accuracy and a large field of view, which can fully meet the robot's needs for motion planning in a large working space. However, after analysis, in unstructured working scenes, the field of view of a single depth camera is affected by obstacle occlusion, so it is necessary to introduce multiple depth cameras to form a high real-time global vision system to obtain more complete environmental information and improve the success rate and safety of

motion planning. Therefore, efficient modelling of unstructured environments is one of the main research directions of this paper.

3. Predictive Control for Robot Vision Feature Completeness Analysis and Global Calibration Analysis

3.1. Predictive Control for Robot Vision Feature Completeness Design. Visual servo tasks require optimization of the overall performance of the task from a macroscopic perspective and the achievement of planned control goals at each step from a microscopic perspective. Both macroscopic optimization and microscopic control require description and feedback of the task in terms of visual features [16]. The completeness of the information contained in the macro- and microfeatures affects the completability of the task. Therefore, the complete set of visual features for a visual servo task is defined to describe the set of features needed to satisfy the task completion, which consists of a complete macrofeature set and a complete microfeature set. Macrofeatures are often built on top of microfeatures, so this section discusses the complete microfeature set first. Robot vision servo control is an interdisciplinary research field, involving multiple disciplinary elements such as image processing technology, computer vision, control theory, and real-time computing. However, the purpose of its task execution is to achieve control goals in the task space. Microscopic visual features play a role in visual servo control to characterize and feedback the control condition of the system. Therefore, evaluating the merit of a microscopic visual feature set for a visual servo task examines whether the hand-eye relationship adequately links up control and visual features. That is, the microscopic visual features should be sufficiently informative to ensure that every control degree of freedom is characterized and fed back by the visual features.

$$f = \frac{\partial H(q)}{\partial q}\dot{q},$$
$$f = J \cdot U^2. \tag{1}$$

The image Jacobian matrix varies with robot poses, so for a particular set of microscopic features, it is important to ensure that the image Jacobian matrix is a full rank within the robot motion range involved in the task. Based on ensuring that the visual servo task can be completed, the feature set with the best properties is selected so that the task completion effect is optimized. First, the mapping relationship between the image space and the task space should be as stable as possible, changing substantially as the relative poses between the target object and the camera change so that the feature set can avoid the system converging to local minima or having unintended motions in the task space [17]. Second, the rate and magnitude of change of the microscopic visual features should be as consistent as possible with the control volume to better characterize it. The microscopic features are chosen to avoid singular values and local

minima as much as possible. When singular values or local minima are present, both visual feature changes and control volume changes will be extremely inconsistent. At singularities, it is known from the hand-eye mapping relationship that a nonzero velocity vector at the end of the actuator corresponds to a visual feature deviation that may be zero, and thus, the visual feature does not correctly reflect the condition of the controlled robot. At local minima, the feature deviations belong to the zero space of the generalized inverse matrix of the image Jacobian matrix.

$$f + f^* = \operatorname{Ker} J^+. \tag{2}$$

The visual feature bias is not zero, and the robot control speed command is calculated to have a value of zero, so the visual feature cannot be used to guide the control either. In addition, good decoupling of the hand-eye relationship facilitates control and planning to avoid control of individual degrees of freedom causing undesired operations in other degrees of freedom, resulting in reduced control efficiency or actuator end motion out of the field of view, as shown in Figure 1.

For the human-machine collaborative operation system, there is a state that the human-machine is not in the same space at the same time; then the robot's working space is an unstructured environment occupied by static obstacles. Unlike the free space or structured environment where the state is known, the work objects and obstacles in the unstructured environment are irregular and randomly placed, especially for the customized and small-batch operations where the human-robot collaborative production method is applicable; the work objects are of many types and often change production. The navigation system is the smallest complete system that integrates environment modelling, scene understanding and reasoning, autonomous decision-making and execution, and human-computer interaction. It corresponds to a complete processing flow of information expression, association, fusion, reasoning, and management. In the actual operation process, the robot first needs to identify the target workpiece among multiple types of work objects, that is, to achieve the position estimation of the target workpiece, which is the basis for completing the subsequent operation tasks and is exactly what was studied in the previous section of this paper [18]. In the process of executing subsequent tasks, the robot calculates the corresponding end tool posture based on the workpiece posture and moves safely and autonomously to the target posture in

an unstructured environment, which is an important manifestation of the robot's intelligence capability, and motion planning is one of the key technologies to realize the robot's autonomous motion, which is also the focus of this chapter. Since robots for human-robot collaboration usually have six or seven degrees of freedom, their planning space is a complex high-dimensional space, which leads to complex space modelling and increased computational effort for collision detection, affecting the efficiency and success rate of motion planning. To solve the above problems and quickly plan feasible motion paths, this chapter first introduces the planning principle of the classical fast extended random tree algorithm and then proposes a heuristic-guided fast extended random tree algorithm for its problems of randomly scattered sampling points and slow convergence, which uses a guided search strategy to accelerate the convergence of the algorithm, and heuristic functions are used to optimize the motion paths to make the optimal motion path length.

$$J = \begin{bmatrix} \dfrac{\lambda_1}{z_1} & 1 & -\dfrac{f_1}{z_1} \\[2ex] 1 & \dfrac{\lambda_1}{z_1} & -\dfrac{f_2}{z_1} \\[2ex] \dfrac{\lambda_1}{z_1} & 1 & -\dfrac{f_3}{z_1} \end{bmatrix}. \tag{3}$$

The camera of a specific focal length can be selected according to the range of features in the actual experiment to avoid singular Jacobian matrices when the inverse matrix zero space of the image Jacobian matrix contains only zero vectors and avoids local minima. Thus, the feature set consisting of the pixel coordinates of the marker points is the complete feature set for this task. With the feature set complete, its performance is discussed. The image Jacobian matrix does not change rapidly with the relative positional change of the hand-eye, so the linear nature is acceptable. Good consistency of hand-eye variation is obtained by avoiding singular values and local minima with a suitable choice of camera focal length. There are three zero terms in the image Jacobian matrix, so the hand-eye relationship under this feature set is well decoupled. The simplicity of the points makes the feature set acquisition efficient and satisfies the requirement of processing speed in human-computer interaction.

$$\theta_0 = \arctan\left(\frac{2\left((M_{11}/M_{00}) + x_\lambda y_\lambda\right)}{\left((M_{11}/M_{00}) + x_\lambda^2 y_\lambda^2\right) + \left((M_{11}/M_{00}) - x_\lambda y_\lambda\right)}\right), \tag{4}$$

$$f(k-1) = f(k) + J(q(k)) \cdot \Delta q\left(k_k^2\right).$$

The autonomous navigation behaviour of a mobile robot is first based on stereo vision sensors to obtain information about the surrounding environment and construct an

environmental possibility map and plan a reasonable executable path under the a priori information of the possibility map, and finally, the robot follows this executable path to
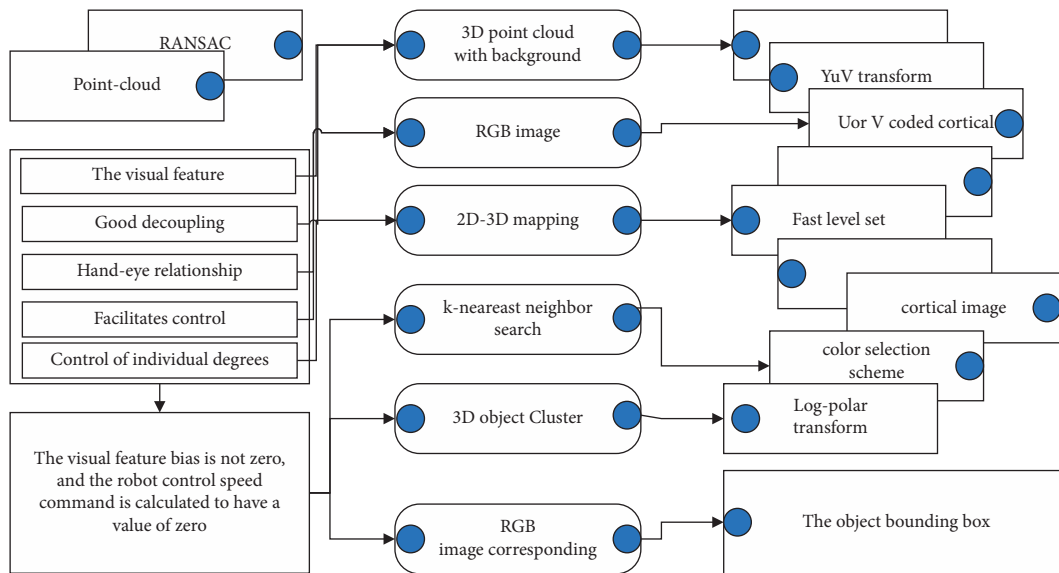
FIGURE 1: Robot vision feature completeness framework.

complete the obstacle avoidance navigation task [19]. The key to the robot navigation problem is to be able to find a safe and smooth path in the process of moving towards the target point while enabling the mobile robot to accomplish the task of dynamic obstacle avoidance. The existing mobile robots have a certain autonomous navigation capability, but because the characteristics of the field environment are more complex and dynamic, the control error and measurement error of the mobile robot is greater; therefore, it is necessary to conduct further in-depth research on the key technology motion planning of the autonomous navigation behaviour of the mobile robot in the field environment, as shown in Figure 2.

For the visual tracking constraint, since this kind of constraint is performed between the current frame and the previous keyframe with a short visual tracking baseline, it can be assumed that the relative motion between the two images is small, and it is feasible to use the optical flow method for data matching alignment in that case. Three indicators are proposed to evaluate the performance of a complete microfeature set. Taking the task of a humanoid robot grabbing a brush as an example, the acquisition and application of a complete set of microfeatures in the task of robot visual serving are studied. Also, it is appropriate to use the optical flow-based method to handle the visual tracking constraint calculation with a higher establishment frequency because of the simple and efficient data matching process. Based on the matching results of the optical flow method in the keyframe selection stage, pixel matching between two frames can be simply obtained to complete the data association establishment of visual tracking constraints.

To avoid the computationally expensive problem of too many candidate regions due to probabilistic sampling, the sampling mode of other traditional tracking algorithms is replaced by a local search cantered on the location of the target, which reduces the search range and improves computational efficiency. The advantage of local search is

that it considers the continuity of the target's motion trajectory and the smoothness of the appearance transformation during the tracking process. The previous frame filter model, obtained by using the filter template update method, is correlated with the local search region of the current image frame to locate the position of the tracking target in the current frame.

$$y = F^{-1}\{Z \oplus F^*\}. \tag{5}$$

To introduce multiscale estimation in the algorithm to obtain an accurate extraction of the region where the target is located, but not to lose the rapidity of tracking efficiency due to the increase in computational complexity, an incremental estimation strategy is used. The strategy separates the position estimation from the scale estimation, forming a cascading relationship. After relying on position estimation to capture the location of the target from the image, scale estimation is enabled to extract the extent of the target's region in the image. This estimation strategy does not affect the computational process of location estimation but also relies on its results to reduce the scale estimation range and reduce the computational cost.

3.2. Experimental Analysis of Visual Feature Completeness Bureau Calibration. Based on the positional information of the operation object, the robot needs to plan a feasible path in the unstructured environment. Since in unstructured scenarios such as human-robot collaboration, there is a situation where the human and the robot are in the same workspace at the same time when the robot is in a dynamic unstructured scene, the state changes of the human, the robot, and the environment are random, so the robot needs to intelligently sense the state changes of the obstacles in the workspace through vision sensors, and when the obstacles may interfere with its normal operation, the robot needs to react in time, that is, plan its motion trajectory online to
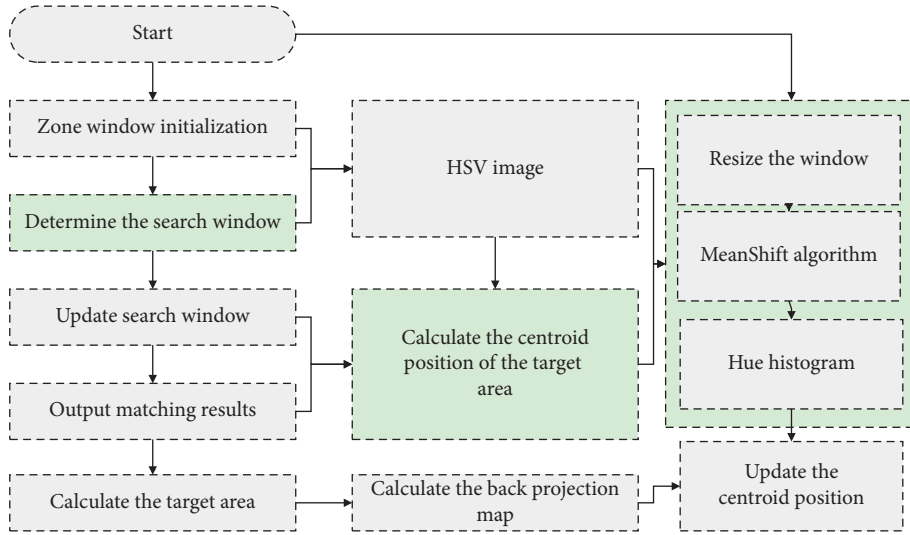
FIGURE 2: Predictive control steps.

achieve obstacle avoidance to avoid potential collision hazards [20]. Therefore, this chapter proposes an online real-time trajectory planning method for robots in dynamic unstructured environments.

First, to solve the problems of the limited single-camera field of view and real-time dynamic obstacle avoidance, a visual perception method is proposed to establish an offline mapping and online fusion model of multicamera depth images and robot workspace 3D raster, determine the occupation state of the occluded 3D raster by the robot and obstacles, and obtain the nearest distance and relative velocity between the stable robot and obstacles in real-time based on Kalman filtering algorithm. It often does not consider the generation of commands for movement control but only considers how to achieve environmental reconstruction and real-time positioning under given information, which is essentially a real-time version of the field of computer vision. Then, based on the improved artificial potential field method to calculate the attractive and repulsive forces, the trajectory avoidance strategy is adjusted according to the relative position and velocity of the obstacles, and the potential field forces are converted into robot joint velocities to control the robot from the velocity level to complete the obstacle avoidance task so that the robot can avoid obstacles and ensure the operational efficiency at the same time. Finally, the effectiveness of the trajectory avoidance algorithm is verified in scenarios with different dynamic obstacles, as shown in Figure 3.

Unlike current path planning methods that optimize the path search approach to improve efficiency, this chapter improves the efficiency of path search by reconstructing the environment representation model. The multilevel graph model-based environment representation method can describe the multilevel topological structure information of the environment more effectively, which in turn improves the efficiency of using environment-inspired information in the path search process. This experiment further validates the idea that building a task-oriented environment
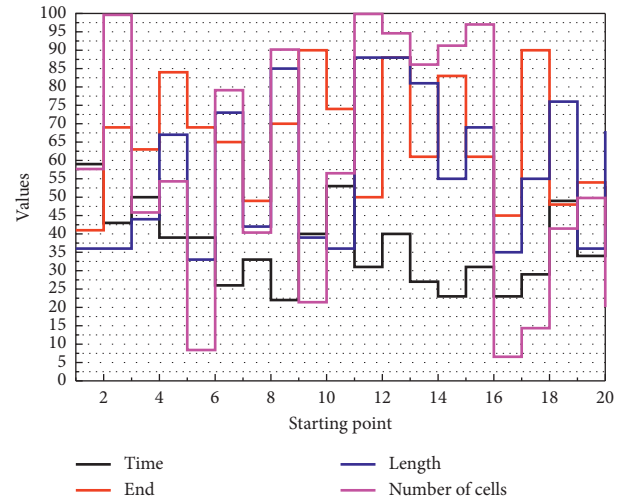
FIGURE 3: Path planning.

representation model can improve the efficiency of retrieving and utilizing information during task execution. However, in the path tracking process, to track the feature paths, the mobile robot needs to accurately identify the global features of the environment, thus requiring higher performance of global localization. The mobile robot needs to globally localize not only its pose but also the features of the environment.

Based on the above description, the definition of a visual tracking failure of the system is given here: a keyframe with an empty set of constraints is considered to have a tracking failure at that point. When a tracking failure occurs, the system considers this keyframe as the first keyframe in a new subgraph at the back end and reestablishes the subgraph coordinate system based on this frame so that future visual tracking is maintained in this new subgraph. It is important to note that in contrast to the first keyframe initialized by the system that is cured and not involved in the optimization

adjustment, the first node of the newly created subgraph here is not cured and will be involved in the future global graph optimization adjustment process. With this subgraph mechanism, the system will always be in two states: maintaining visual tracking in a subgraph or creating a new subgraph and trying to track maintenance in that subgraph. Thus, the system can run all the time without running out. The flow diagram of the back end of this subgraph is shown in Figure 4.

During the visual tracking of a particular subgraph, the poses of the nodes are given according to the motion model. The pose estimate of the current image frame is determined by the known pose of the related frames within the same subgraph and the constraints between the two frames. Also, nodes that are within the same subgraph will be influenced by other nodes that are constrained by another subgraph, thus achieving an overall alignment of the subgraph. Mobile robots are the most important category of robots. With the continuous development of human society and the continuous improvement of informatization and intelligence, mobile robots will surely replace human work partially or completely in industries such as manufacturing industry, agriculture, national defense, and scientific research. Thus, visual tracking always operates within a certain subgraph, and the operation of that visual tracking is always a two-view geometric operation based on two frames of images. With such a strategy, it is possible to maintain the operation of the system in the face of visual tracking failure in this system, despite the loss of information about other existing nodes.

When building a workstation for human-robot collaboration, the area of interest to be observed needs to be set and multiple cameras need to be arranged according to the scope of the working scene regarding the above method to obtain more complete environmental information [21]. Although two cameras can acquire more complete environmental information, the amount of data acquired by two depth cameras at the same time is large, and they need to be updated and fused in real time, which requires more computational effort and time, resulting in the robot's intelligent perception and trajectory planning performance in dynamic unstructured environments being affected by the online updating efficiency of the environmental model. Therefore, to speed up the fusion and update of dual-camera data, this paper draws on the idea of offline modelling and online updating, builds a model of the projection relationship between 3D raster and depth image in the offline stage based on the projection principle of 3D space and 2D image, calculates the projected pixel points of each 3D raster centre in the robot workspace on the depth image, and constructs the mapping data structure of 3D raster to depth map pixels to the data set that is saved offline.

4. Results and Analysis

4.1. Predictive Control Results for Completeness of Robot Vision Features. Figure 5 depicts the entire process of right-hand tracking for the NAO robot. In this task, the binocular vision system takes pictures of the workspace and the vision servo system uses the Camshafts algorithm to detect the end of the actuator and the quill. In this way, this process of the task can be described in vision space. At the beginning of the task, the NAO robot's left and right cameras capture images with a difference of 111 pixels in the horizontal and 163 pixels in the vertical direction. A Kalman–Bucy filter-based visual servo controller is used to estimate the image Jacobian matrix, which in turn derives control quantities to act on the robot. The NAO robot eventually reaches the quill within 17 steps, and the deviation between the two centres is within 10 pixels. The experimental results show the robustness of the algorithm proposed in this paper to image noise, motion disturbances, and irregular motion of the target object.

Task-oriented evaluation of visual feature sets examines the impact of feature sets on task completability and completion performance. First, a definition of a complete feature set for a visual servo task is proposed, and a complete feature set includes a complete microfeature set and a complete macrofeature set. This section focuses on the complete microfeature set and proposes a criterion to determine whether a feature set is a complete microfeature set for a visual servo task. On this basis, if the visual feature set is complete, its performance needs to be further examined in terms of the linear nature, change consistency, and decoupling of the complete feature set. Taking a humanoid robot grasping a brush as an example, the acquisition and use of the complete microscopic feature set is investigated, and a visual servo system is proposed, which is built based on the calibration-free Kalman–Bucy filter visual servo method and the Camshaft detection algorithm. Using this system, the NAO robot accomplished the grasping of brushes well in the calligraphy task. The experimental results verify the validity of the theory of complete microscopic feature sets proposed in this chapter and the robustness of the complete feature set extraction algorithm to image noise, humanoid robot return error, and an irregular motion of the target object.

In terms of quantitative evaluation for the multisugar back end, to demonstrate the improved enhancement of this back end design for fault recovery, we recorded the timestamps of the occurrence of tracking loss and closed-loop detection during the experiment, as well as the localization error at the corresponding time points. The results are shown in, Figure 6, where the difference between the true bit poses recorded by the red grasping system is obtained, the light blue line is the response marker for the moment of tracking loss, and the red line is the response marker for the occurrence of closed loop. During the figure, visual tracking, as well as map construction, can always maintain operation in either subgraph, and when mutual constraints are detected between the subgraphs, the error curve decreases due to the alignment between the subgraphs. Furthermore, the trend of the error curve shows that after a certain time of environmental exploration, the error curve stops rising, and any new keyframe always finds a relevant reference frame in the map, which in turn establishes a constraint and achieves successful localization. Such curve results quantify the ability of the system in tracking lost fault recovery.
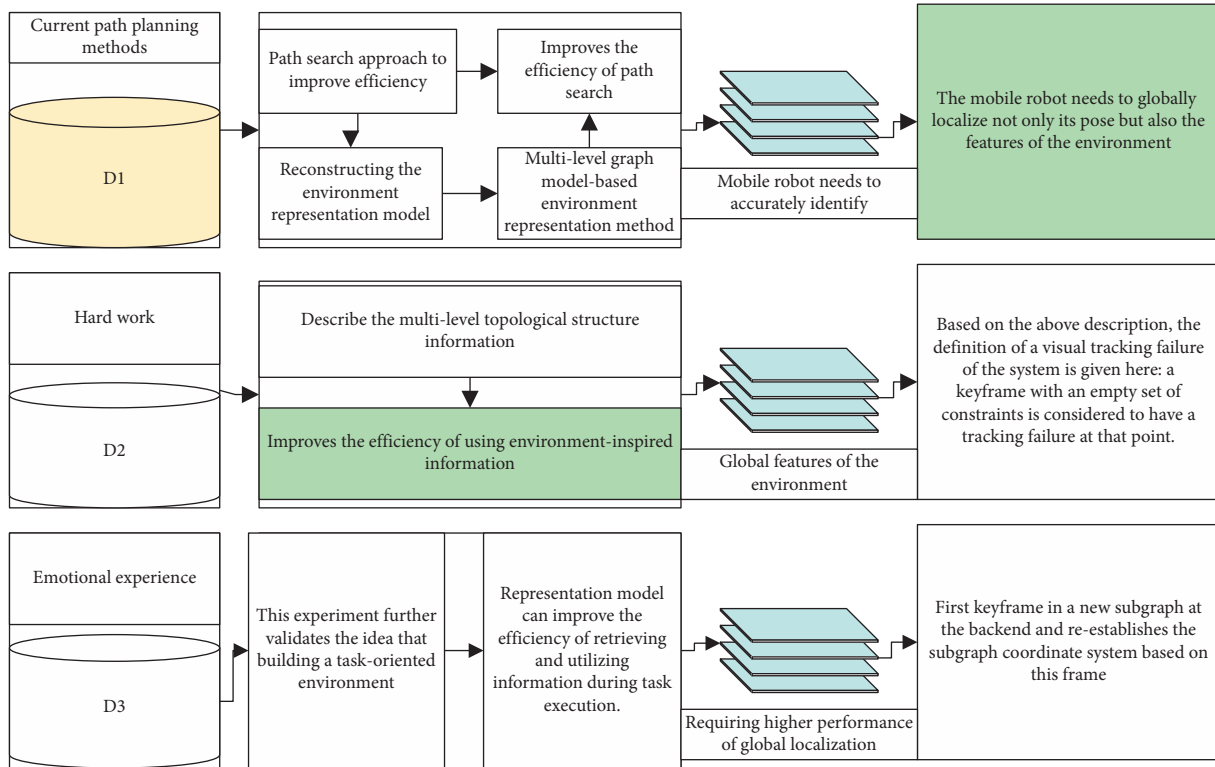
Figure 4: Back-end flow diagram based on multiple subgraphs.
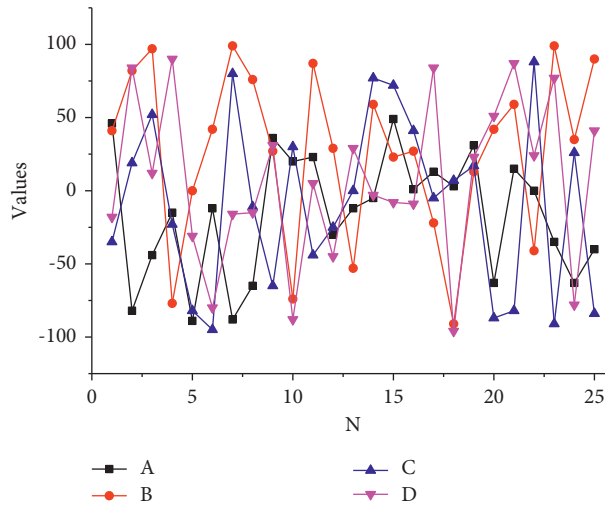
Figure 5: Observed tracking bias.

The decreasing trend in the number of particles employed in the process indicates that the robot positional uncertainty is reducing the influence of its fluctuation process by the environmental structure. In the algorithm of this chapter, the number of effective particles is always maintained at a low level, which makes the algorithm require a low amount of computation. As the mobile robot moves, it collects enough information about its environment, and the uncertainty in its positional estimation decreases. However, to still ensure the robustness of the positional estimation, the algorithm still maintains a certain number of particles for the estimation of the positional pose.

The convergence rate is low in the corridor scenario because only the boundary information of the motion model is provided and the inner content is not fully described. Also, due to the use of omnidirectional observation particles, while obtaining as many possibility regions as possible, it introduces correlation uncertainty that needs to be fused with more observations for elimination. From the comparison experiments, the initial likelihood region calculation has an important impact on the convergence speed of the global localization algorithm. In the corridor scenario, it is difficult to make an accurate estimate of the region in which the initial poses are located. However, in the results of the

FIGURE 6: Analysis of positioning error accuracy at the back end of multisubgraph.

method in this chapter, despite the large deviation of the initial moment's positional estimation, the algorithm is still able to converge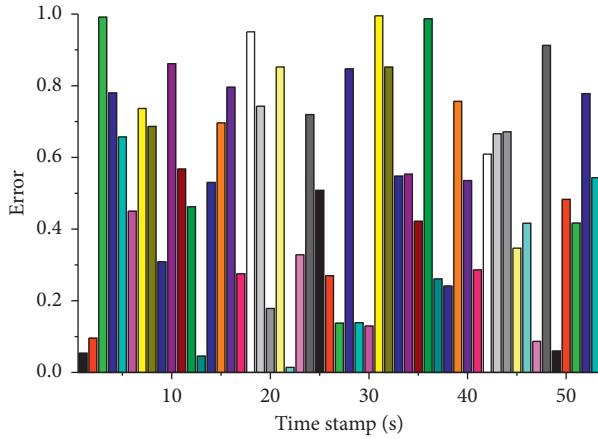 to the correct region after a period of motion, which fully verifies the robustness and stability of the method in this chapter in the process of global localization. It compensates for the information fed back from the system, relaxes the requirements for the mechanical accuracy and joint tightness of the robot components so that the robot system can also improve flexibility in an uncertain external environment, and realizes various automatic control functions. Through simulation experiments, this chapter verifies the effectiveness of the proposed method in terms of global localization accuracy and further verifies the effectiveness of this chapter's method in terms of global localization efficiency and global localization robustness through online experiments.

4.2. Experimental Results of Visual Feature Completeness Bureau Calibration. Geometric features in scenes are pervasive features, less affected by texture, more robust to changes in perspective, and so on. Geometric features can improve the ability to represent space. Establishing a mapping model of geometric feature descriptors to their spatial locations and constructing a method to retrieve their corresponding geometric features for a given spatial location are essential to improve the ability of geometric features to represent space. There are two main types of geometric features. One is the features corresponding to geometric primitives, such as line segments, planes, and other features. The second is local geometric features of dense surfaces, which do not have a specific form of expression but describe geometric information of the corresponding surface and can generally be modelled using surface elements. Single features have limited ability to characterize spatial information and are more dependent on environmental conditions. Exploring global localization methods for multiple features is important for improving navigation system stability and environmental adaptability.

Line segment features are pervasive features in the environment and are the main environmental features in

weakly textured scenes. Their number is generally small, but the feature descriptors are relatively expressive. In this chapter, a visual global localization method based on multivariate geometric features is proposed, which can better reason about feature associations by using the relationship between multivariate features. Furthermore, since the dense geometric information of the environment can greatly improve the screening quality of the positional hypothesis, this chapter explores more efficient dense map representation methods that enable the robot to quickly retrieve location-specific geometric distribution information to speed up the screening efficiency. A schematic diagram of visual relocation based on multivariate geometric features is shown in Figure 7.

In the interference observer-based visual servo control, the tracking results observed from both cameras are shown in Figure 8. It can be seen from the figure that the system can track the target object well using the interference observer-based algorithm even in the presence of image noise and external interference. It can be concluded that the visual servo method based on interference observer can suppress the effect of image noise and external interference well while ensuring the closed-loop stability of the system. At the same time, there is no steady-state error.

The experiments were first conducted using a pioneer robot equipped with vision sensors to collect training data and test data in a laboratory scene and a corridor scene. The lab scene contains richer texture information, while the corridor scene is a typical weak texture scene. Line segment features are the main feature type in the corridor. The data sequences collected in the static laboratory scene are Official and Office 2. To test the robustness of the method in this chapter to dynamic objects, pedestrians are set to walk during the test, and the corresponding data sequences are Office-D and Corridor-D. The method in this chapter is compared with the DBoW2 method, which is a global localization method widely used in current SLAM systems. It can be seen from Figure 8 that the method in this chapter can obtain more successful localization frames, in which, the indexes of the training and test sequences are given in the table. The experimental indexes are the number of frames with localization error less than 5 cm and the number of frames less than 10 cm. In the laboratory scenario, more correct localization frames can be obtained due to the richer features. Comparing the localization results of ORB-SLAM2 in Office and Office-D, the robustness of the method in this chapter is higher for dynamic objects. Although the method in this chapter obtained fewer successful localization frames with errors less than 5 cm in the corridor scene, most of the frames were localized with errors of 10 cm, indicating the robustness of the method in this chapter for dynamic, low-texture scenes.

The feature point maps, line segment feature maps, and dense point cloud maps learned through regression trees, which greatly improve the efficiency of feature-based retrieval of space and spatial information-based retrieval of features. In the process of global localization, the camera poses are estimated by combining the applied point information and line segment information, and stable point
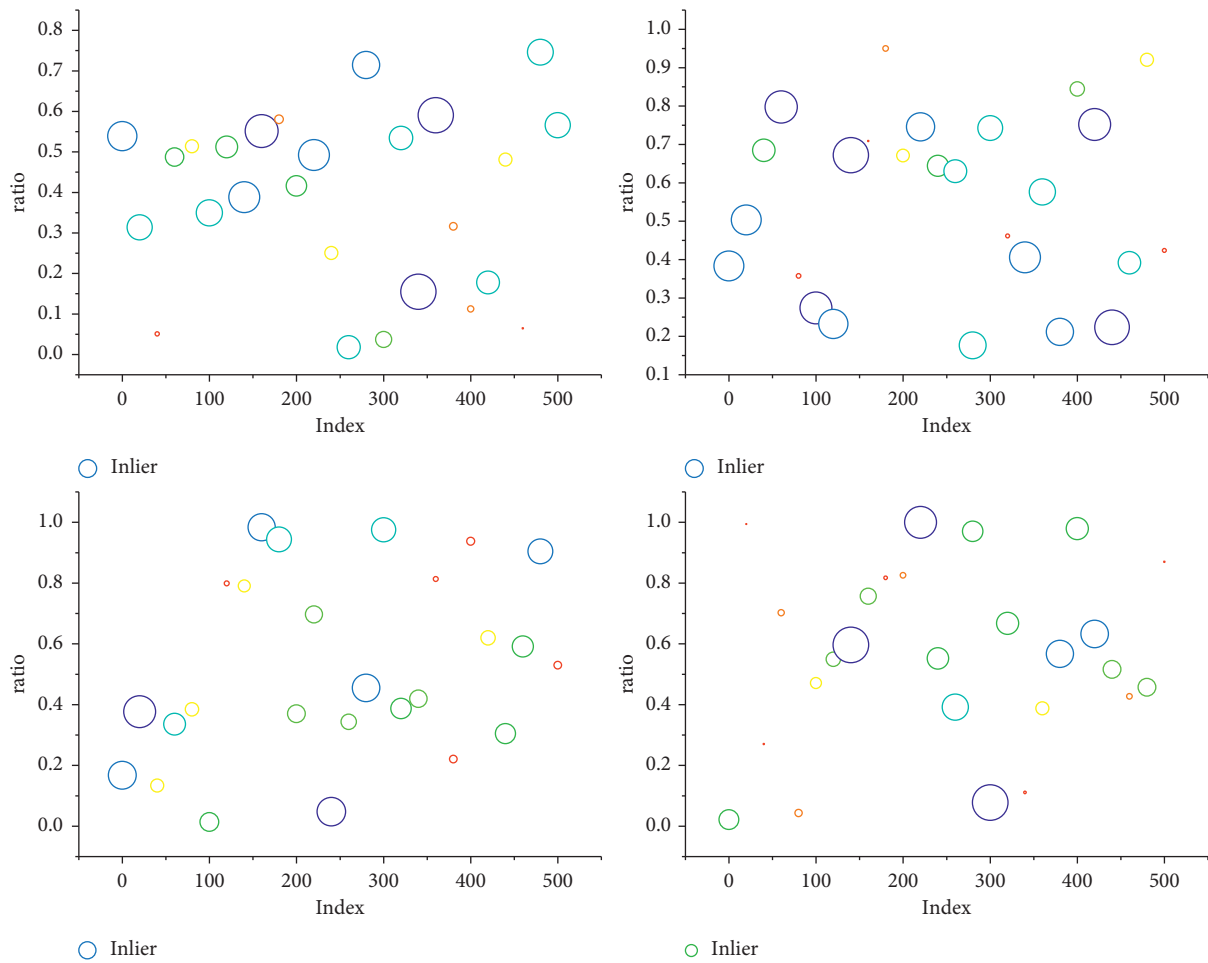
FIGURE 7: Results of point matching within the test image.
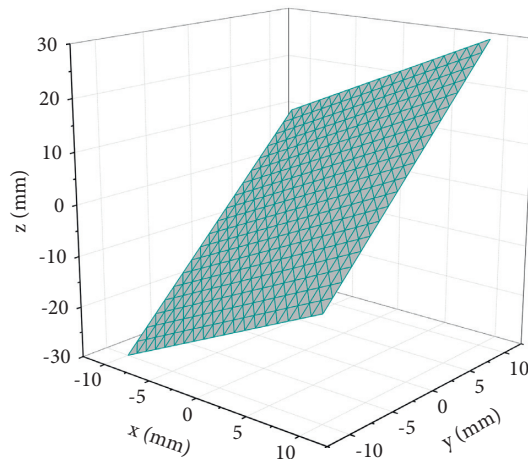
FIGURE 8: Tracking trajectory obtained using interference observer-based vision in a workspace.

pair matching is constructed using point-to-linear projection to realize the computation of positional assumptions. The stacked RANSAC method is constructed to address the problem of insufficient sampling efficiency due to the low proportion of interior points in the matched pairs. Compared to the traditional RANSAC method, the stacked RANSAC method uses the results of the previous sampling information to guide the new sampling process to

be more biased towards meaningful samples. At the same time, a more efficient method of assuming a quality metric for the poses is utilized to constrain the current observation to match the environmental surface in a more efficient manner using the dense geometric information of the environment.

5. Conclusion

The concept of a complete feature set is proposed from the perspective of visual servo task completability; the determination method of a complete microscopic feature set is given; and the extraction and use of a complete microscopic feature set are investigated using a humanoid robot grasping a brush as an example. In this paper, the hand-eye relationship of the feature-based 2D visual servo task is described by the image Jacobian matrix, and the completeness of the microscopic feature set is examined by finding the rank of the image Jacobian matrix to see whether the microscopic feature set completely characterizes the robot control degrees of freedom, that is, the completeness of the microscopic feature set is examined. The camshaft method is used to track the target object, obtain the microscopic feature set, and complete the task based on the complete microscopic feature set using a Kalman–Bucy filter-based visual serving method. Through the experiments based on the NAO robot, it is confirmed that the feature set proposed in this paper can guarantee the completion of the task, the microscopic feature set acquisition method has good robustness in the presence of image noise and irregular motion trajectory of the target object, and finally, the visual serving task can be completed with good performance according to the whole visual serving strategy. Based on the Q-filter to construct the interference observer, the equivalent interference consisting of model uncertainty, high-frequency detection noise, and low-frequency input interference is estimated and eliminated through the input and output signals of the controlled object so that the part consisting of the controlled object and the interference observer externally exhibits the given nominal model, and then the visual servo controller is designed according to the nominal model to achieve robustness under the premise of closed-loop stability of the system high-performance control effect under the premise of closed-loop stability of the system. Finally, the effectiveness of the calibration-free visual servo method proposed in this paper is verified by comparing the algorithm with the proportional-integral method and the divisional Brayden method in the simulation environment and comparing the method with the proportional-integral method in the experimental conditions.

Data Availability

The data used to support the findings of this study are available upon request to the author.

Conflicts of Interest

The author declares that there are no conflicts of interest in this article.

References

[1] C. Zhang, "Binocular vision navigation method of marine garbage cleaning robot in unknown dynamic scene," *Journal of Coastal Research*, vol. 103, no. SI, pp. 864–867, 2020.

[2] M. Dirik, A. F. Kocamaz, and O. Castillo, "Global path planning and path-following for wheeled mobile robot using a novel control structure based on a vision sensor," *International Journal of Fuzzy Systems*, vol. 22, no. 6, pp. 1880–1891, 2020.

[3] H. Liu, D. Qu, F. Xu, F. Zou, J. Song, and K. Jia, "Approach for accurate calibration of RGB-D cameras using spheres," *Optics Express*, vol. 28, no. 13, pp. 19058–19073, 2020.

[4] Q. Liang and M. Liu, "A tightly coupled VLC-inertial localization system by EKF," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3129–3136, 2020.

[5] L. Lattanzi, C. Cristalli, D. Massa, S. Boria, P. Lépine, and M. Pellicciari, "Geometrical calibration of a 6-axis robotic arm for high accuracy manufacturing task," *International Journal of Advanced Manufacturing Technology*, vol. 111, no. 7, pp. 1813–1829, 2020.

[6] H. Alzarok, S. Fletcher, and A. P. Longstaff, "Survey of the current practices and challenges for vision systems in industrial robotic grasping and assembly applications," *Advances in Industrial Engineering and Management*, vol. 9, no. 1, pp. 19–30, 2020.

[7] S. Raileanu, T. Borangiu, A. Silisteanu, S. Anton, and F. Anton, "Open source platform for vision guided robotic systems integrated in manufacturing," *Journal of Control Engineering and Applied Informatics*, vol. 22, no. 4, pp. 52–60, 2020.

[8] C. Park, P. Moghadam, S. Kim, S. Sridharan, and C. Fookes, "Spatiotemporal camera-LiDAR calibration: a targetless and structureless approach," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1556–1563, 2020.

[9] P. Bauer, F. Fink, A. Magaña, and G. Reinhart, "Spatial interactive projections in robot-based inspection systems," *International Journal of Advanced Manufacturing Technology*, vol. 107, no. 5, pp. 2889–2900, 2020.

[10] D. Tish, N. King, and N. Cote, "Highly accessible platform technologies for vision-guided, closed-loop robotic assembly of unitized enclosure systems," *Construction Robotics*, vol. 4, no. 1, pp. 19–29, 2020.

[11] E. Bjørne, E. F. Brekke, T. H. Bryne, J. Delaune, and T. A. Johansen, "Globally stable velocity estimation using normalized velocity measurement," *The International Journal of Robotics Research*, vol. 39, no. 1, pp. 143–157, 2020.

[12] L. Yang, I. Dryanovski, R. G. Valenti, G. Wolberg, and J. Xiao, "RGB-D camera calibration and trajectory estimation for indoor mapping," *Autonomous Robots*, vol. 44, no. 8, pp. 1485–1503, 2020.

[13] M. Shao and M. Hu, "Parallel feature based calibration method for a trinocular vision sensor," *Optics Express*, vol. 28, no. 14, pp. 20573–20586, 2020.

[14] S. Assadzadeh, C. K. Walker, L. S. McDonald, P. Maharjan, and J. F. Panozzo, "Multi-task deep learning of near infrared spectra for improved grain quality trait predictions," *Journal of Near Infrared Spectroscopy*, vol. 28, no. 5-6, pp. 275–286, 2020.

[15] Y. Zhao and P. A. Vela, "Good feature matching: toward accurate, robust vo/vslam with low latency," *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 657–675, 2020.

[16] Y. Jiao, Y. Wang, X. Ding, B. Fu, S. Huang, and R. Xiong, "2-Entity random sample consensus for robust visual localization: framework, methods, and verifications," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 5, pp. 4519–4528, 2020.

[17] F. Yan, Z. Liu, X. Pan, and Y. Shen, "High-accuracy calibration of cameras without depth of field and target size limitations," *Optics Express*, vol. 28, no. 19, pp. 27443–27458, 2020.

[18] Q. Liu, X. Xie, X. Zhang et al., "Stepwise calibration of plenoptic cameras based on corner features of raw images," *Applied Optics*, vol. 59, no. 14, pp. 4209–4219, 2020.

[19] M. A. Rafique and A. F. Lynch, "Output-feedback image-based visual servoing for multirotor unmanned aerial vehicle line following," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 4, pp. 3182–3196, 2020.

[20] I. Kholodilin, Y. Li, and Q. Wang, "Omnidirectional vision system with laser illumination in a flexible configuration and its calibration by one single snapshot," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 9105–9118, 2020.

[21] H.-W. Chae, J.-H. Choi, and J.-B. Song, "Robust and autonomous stereo visual-inertial navigation for non-holonomic mobile robots," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9613–9623, 2020.

Research Article

Research on the Application of GIS Technology Combined with RBFNN-GA Algorithm in the Delineation of Geological Hazard Prone Areas

Tianwang Lei [ID],[1] **Yao Lu** [ID],[1] **Chong Zhang** [ID],[2] **Jing Wang** [ID],[1] **and Qi Zhou** [ID][2]

[1]*School of Civil Engineering, Xi'an Traffic Engineering Institute, Xi'an, Shannxi 710300, China*
[2]*School of Geography and Environment, Baoji College of Arts and Sciences, Baoji, Shannxi 721013, China*

Correspondence should be addressed to Chong Zhang; zhangchong@bjwlxy.edu.cn

With the rapid development of the economy and society, geological disasters such as landslides, collapses, and mudslides have shown an intensifying trend, seriously endangering the safety of people's lives and property, and affecting the sustainable development of the economy and society. Aiming at the problems of merging different data layers and determining the weighting of data stacking in the statistical analysis model based on GIS technology in the evaluation of the risk of geological disasters, this study proposes a logistic regression model combined with the RBFNN-GA algorithm, that is, the determination of the occurrence of geological disasters. The fusion coefficient (CF value) with the RBFNN-GA algorithm model, and with the help of SPSS statistical analysis software, solves the problem of factor selection, heterogeneous data merging, and weighting of each data layer in the risk assessment. In the experimental stage, this study adopts the method of geological hazard certainty coefficients to carry out the sensitivity analysis of the geological hazards in the study area. Using homogeneous grid division, the spatial quantitative evaluation of the risk of geological disasters is realized, and at the same time, the results of the spatial quantitative evaluation of the risk of geological disasters are tested according to the latest landslide points in the region. The existing classification mainly depends on the acquisition of land use/cover information or the processing method of the acquired information, but the existing information acquisition will be limited by time, space, and spectral resolution. The results show that the number of landslide points per unit area in the extremely unstable zone and the unstable zone is 0.0395 points/km^2 and 0.0251 points/km^2, respectively, which is much higher than 0.0038 points/km^2 in the stable zone, indicating the evaluation results and actual landslide conditions.

1. Introduction

According to statistics, in the past 50 years, landslides, collapses, and mudslides have caused more than 20,000 deaths, with hundreds to more than 1,000 deaths every year. Among the many natural disasters, the death toll is second only to earthquakes and floods. In particular, landslides, collapses, and mudslides that erupt in densely populated areas such as towns and mining areas often cause a disaster event that kills hundreds of people [1]. Geological disasters such as landslides, collapses, and mudslides are one of the natural disasters that cause the loss of people's lives and property [2]. In order to effectively prevent and reduce

geological disasters, it is first necessary to have a more comprehensive understanding of regional geological disasters. When formulating regional geological disasters, they can be more targeted. At the same time, research on geological disaster forecasting and early warning is based on the division of geological disasters. The foundation has also become a hot issue in the field of geological disaster research [3]. There are many domestic studies on the mechanism of single geological disasters such as landslides, but the research progress on the development and distribution of regional geological disasters is relatively slow [4]. Exploring and forming a set of scientific, complete, and practical research methods for the development and distribution of regional

geological disasters, and evaluating the risk of regional geological disasters, is of great practical significance for the prevention and control of regional geological disasters [5].

The geological hazard evaluation method adopted at the beginning has inherited the evaluation of geography in many methods, and the methods used can be summarized as follows: dominant factor method, multifactor single-factor item-by-item overlay method, geographic correlation analysis method, and theoretical derivation method [6]. After the 1980s, the degree of quantification increased, the application of mathematical knowledge in geosciences developed vertically and horizontally, and the geological hazard evaluation and zoning research gradually tended to quantify [7]. The extensive application of linear regression prediction, mathematical model analysis, gray theory, neural network, GIS, etc. is gradually introduced to further improve the degree of quantification [8]. In the study of slope geological hazards in an urban development planning project in Australia, Mai [9] integrated the risk, vulnerability, and risk assessment of slope hazards, using GIS software as the technical platform, and using two-dimensional and three-dimensional evaluation systems, respectively, for research on the danger and risk zoning of slope geological disasters in Cairns area. He et al.'s [10] probabilistic analysis and forecasting models show that these models have their own advantages but also have certain shortcomings. For example, the models established by index analysis methods, probability statistical methods, and fuzzy prediction methods are generally only suitable for specific research. At the same time, it is difficult to extract information related to the geology, mechanics, and other environmental conditions of the sliding mass. Liu [11] applied a genetic algorithm to the neural network to form the GA-ANN method to collaboratively solve optimization problems in complex engineering. This method not only uses the functions of neural network's nonlinear mapping, network reasoning, and prediction but also uses the global optimization characteristics of genetic algorithm. It can be widely used in many complex projects where the objective function is difficult to express in the form of explicit function of decision variables. Peng [12] applied the genetic algorithm to the artificial neural network model to determine the maintenance strategy of the parts and analyzed it in combination with examples. After experimental verification, the system selected the maintenance strategy and the actual value is in full compliance with the relative error between the predicted value of the maintenance cost and the actual value. Within the allowable range, the model is credible. Afan [13] applied the genetic algorithm to optimize the parameters of the controller based on the neural network structure and used the controller to control the objects with pure lag. The experiment proved that the control system optimized by the genetic algorithm has good static performance and dynamics. Performance has made a new exploration for solving this problem in the control field. The background factors include engineering rock formations, topography, geological structure, seismic intensity, human engineering activities, and atmospheric precipitation [14–16].

This study takes the regional geological disaster survey project of the autonomous region as the background, based on abundant actual data, deeply researches various factors affecting the occurrence of geological disasters, comprehensively analyzes the relationship between the distribution of geological disasters and various factors, and explores the determination of regional geological disasters. They should have the same or similar spectral information features and spatial information features, so the feature vectors of the pixels of the same type of features will be clustered in a unified feature space area, and different features will be due to different spectral information features or spatial information features. The method of risk has established a risk assessment system for regional geological disasters. Studies have shown that the regional layered soft phyllite intercalated with hard sandstone rock formations, a slope of $20°–40°$, and an elevation range of 320–800 m above sea level are most prone to geological disasters. The key factors affecting the risk of geological disasters are elevation, rock formations, road construction, and slope. According to the calculated cell geological hazard probability value, the regional geological hazard risk is divided into 4 levels, namely, extremely unstable, unstable, basically stable, and stable.

2. GIS Technology Based on RBFNN-GA Algorithm in the Delineation Model Construction of Geological Hazard Prone Areas

2.1. GIS Technology Level Classification. Geographic information system (GIS) is a decision support system that collects, stores, manages, analyzes, and reproduces information related to geographic spatial distribution and has various characteristics of an information system [17]. The main difference between geographic information system and other information systems is that the information stored and processed by it is geocoded. Geographic location and related feature information become an important part of information retrieval. Figure 1 shows the hierarchical topology of GIS technology.

The most important part of the geographic information system is spatial data. Spatial data can effectively express spatial location information and attribute data, while GIS spatial database is a collection of spatial data reasonably stored. The spatial database management system developed on this basis can effectively provide spatial query and analysis [18–20]:

$$p(x(1), x(2), \ldots, x(n)) = \prod_{i=1}^{n} p\{x(n)|x(n-i+1)\}. \quad (1)$$

In MapX, all the features in a layer constitute a feature set. Each graphic element is a feature object (Feature), and many methods of the layer object (Layer) return the feature set of the layer. You can search and locate geographic features such as lines, symbols, or regional characteristics [21–24]. In order to be able to use this method, there must be an index field in the MapInfo table of the search layer:
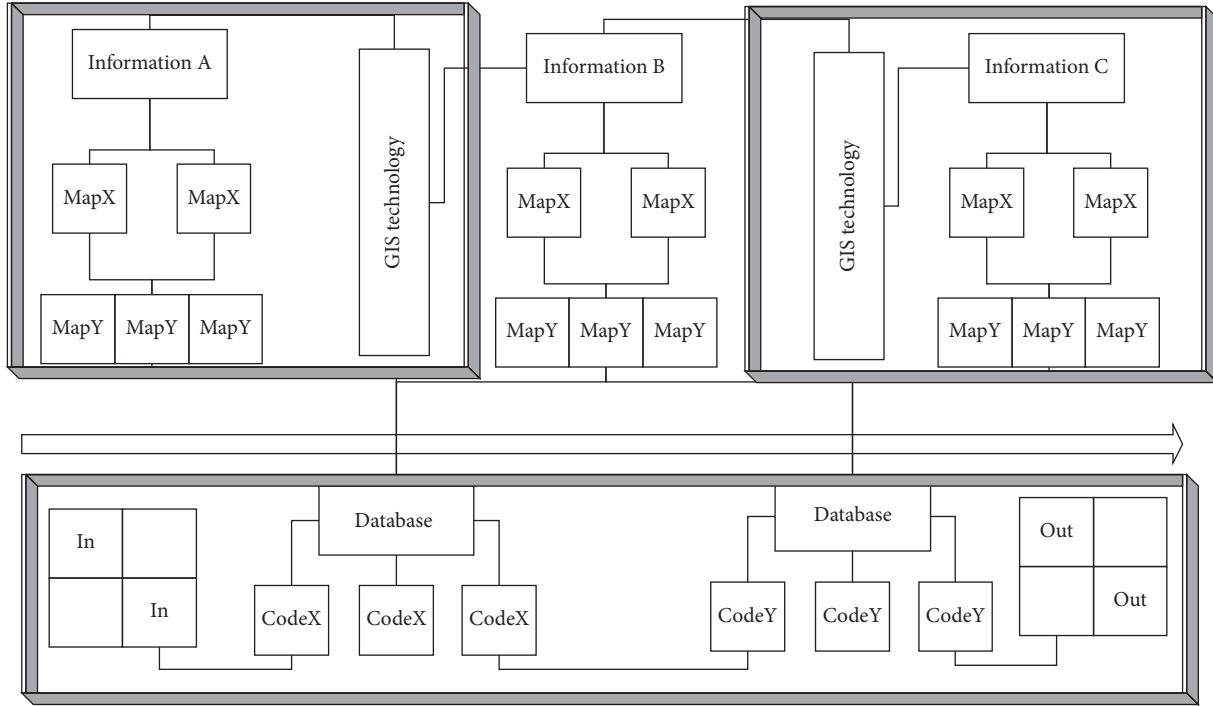
FIGURE 1: Hierarchical topology of GIS technology.

$$f(x) = x(i) + \frac{x(i) \times x(j)}{x(i) + x(j)},$$

$$T^2 - \frac{1}{n} \times \sum_{i=1}^{n} (p(i) - p(x))^2 = 0. \qquad (2)$$

The basic building blocks of MapX components are a single object and a collection. The collection includes objects, which is a combination of multiple objects. Each kind of alignment and collection is responsible for processing a certain aspect of the map. In the model structure of MapX, the Map object exists at the top, and each MapX object, attribute, and method are derived from the Map object.

$$\begin{cases} x(i,0) = a - x(j,0), \\ x(i,1) = b - x(j,1), \\ x(i,j) = \sqrt{\dfrac{x(j,i)}{a+b}}. \end{cases} \qquad (3)$$

Each attribute and method under the Map object will have an impact on the generation of the entire Map object. Each Map object is mainly defined by DataSets, Layers, and Annotations objects. Among them, Layers are mainly used to manipulate the layers of the map, DataSets are used to access spatial data tables, and Annotations are used to add text and symbols on the map:

$$S(x,m) - \frac{(1/2) \times x(m)p(n) + (2/3) \times x(n)p(m)}{p(m) + p(n)} = 0,$$

$$I = \{C(x,y), x \in (0,1,2,\ldots,n), y \in (0,1,2,\ldots,n-1), n \in N\}. \qquad (4)$$

The basic idea of GIS is to divide the main functional modules of GIS into several components, and each component performs different functions. Various GIS components, as well as between GIS components and other non-G1S components, can be easily integrated through visual software development tools to realize the final GIS application [25–27]:

$$L(x,y) = \sqrt{\sum [C_a(x,y) - C_b(x,y)]^2}. \qquad (5)$$

Each map has a data set. Through the data collection, the user's attribute data can be connected with the map spatial data. Data binding is the process of introducing external data into MapX. External data can be multiple types of databases.

$$K(x,y) = \begin{bmatrix} k(x) \times \cos x & 1 \\ 1 & k(y) \times \sin y \end{bmatrix}. \qquad (6)$$

After data binding, you can browse the data on the map as primitives or create a thematic rendering map based on these data. The learning process consists of three stages. The first stage uses unsupervised methods to determine the center of the RBF; the second stage determines the width of the RBF based on the determined center of the RBF; and the third stage determines the distance between the intermediate

layer and the output layer. Generally speaking, these stages are carried out separately.

2.2. The Composition of Geological Hazards. Geological disaster refers to a disaster related to geological action caused by natural factors or human engineering economic activities, which damages the ecological environment and endangers the safety of human life and property. It is in the shape of an inverted stone cone, with a loose and disorderly structure. The dangerous rock mass is a dangerous mountain mass that is cracking and deforming, and may collapse. Landslide refers to the phenomenon that the rock and soil on the slope lose the original balance condition and move downward along a certain weak surface as a whole or scattered along a certain weak surface under the influence of river erosion, groundwater activity (heavy rain), earthquake, and artificial slope cutting. The tendency of the weak structure surface is consistent with the slope direction, and the slope angle is most likely to occur when the slope angle is greater than the inclination angle of the weak surface. The material can be a loose layer, soft rock, and hard rock. A complete landslide should have a sliding surface (slide bed), landslide surface cracks and steps, trailing and lateral sliding walls, cracks, and front bulging. The influencing factors and formation conditions of geological disasters are extremely complex. They act independently, but also influence and overlap each other. According to their active forms, they are summarized as influencing factors and basic conditions; influencing factors include atmospheric rainfall, hydrogeology, neotectonic movement, and human activities; the basic conditions include topography and geomorphology conditions, stratum lithology conditions, and geological structure conditions. The spatial database of this project includes spatial databases of engineering geology and geological disasters. Figure 2 shows the fan chart of the factors of geological hazards.

This database system is constructed in accordance with the relevant provisions. The system includes geological hazard field survey data, result data, its spatial graphic database, and database structure, and the naming of format, layer, view project file, and the structure of primitive number are all carried out in accordance with regulations. The research area has a relatively complete range of disasters. Based on the geological environmental conditions formed in the area, the degree of disaster susceptibility, the degree of harm, and the integrated social development plan, a comprehensive analysis is made. The research area is divided into key prevention and control areas, subkey prevention and control areas, and general prevention area. We adopt measures such as restoring vegetation, cutting slopes, clearing dangerous rocks, strengthening supports, monitoring and forecasting of ground collapse disaster areas, strengthening of the survey of building foundations, assessment of the risk of geological disasters in disaster areas, and dredging of debris flow disaster areas. For river courses, excavation of water-cutting ditches around mountains, construction of slope protection, valleys, blocking dams, etc., the long term should be based on biological engineering and engineering treatment. According to geological survey data, the probability of geological disasters on concave slopes in the
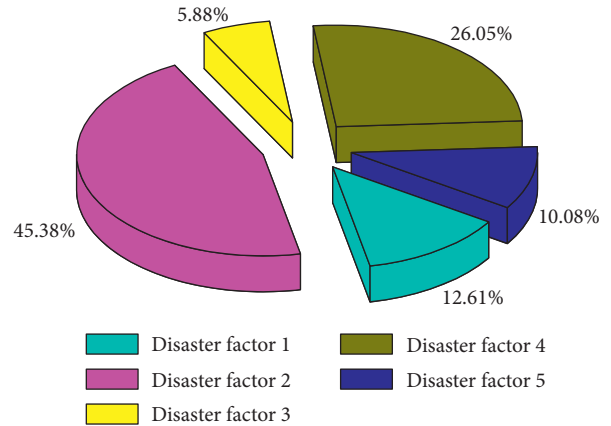

Figure 2: Fan-shaped diagrams of geological hazard factors.

study area is low, and the probability of geological disasters on linear and convex slopes is greater. When the slope of convex slopes in the area is more curved, the relatively concentrated stress of the slope ultimately affects the stability of the slope. Therefore, the shape of the slope directly affects the possibility of geological disasters in the study area.

2.3. RBFNN-GA Algorithm Design. The RBF neural network is a forward neural network that combines a parameterized statistical distribution model and a nonparametric linear perceptron model. The essence of the RBF neural network is the combination of unsupervised clustering method and supervised single-layer linear perceptron to realize the neural network model of nonlinear mapping. The RBF neural network is composed of 3 layers, namely, the input layer, the middle RBF layer, and the output layer. For hyperspectral images, it is difficult to get so many pixels in a limited space for certain classes, and as the dimensionality of the feature space increases, the accuracy degradation called the Hughes phenomenon also requires more training samples. Generally, the number of nodes in the input layer is equal to the dimension of the selected data in the application field, and the number of nodes in the output layer is equal to the number of categories to be classified. For a specific application, the number of nodes in the input layer and output layer is determined. Therefore, the design of the RBF network structure is mainly to determine the number of nodes in the middle RBF layer. It uses the connection strength and the nonlinear input-output relationship of neurons to realize the nonlinear mapping from the input state space to the output state space. Feedforward networks are widely used in pattern classification and feature extraction. Another mode of operation is evolutionary, in which the input is equivalent to the initial state, and the final state of network evolution is the output. This kind of network is similar to a dissipative nonlinear dynamic system.

Figure 3 shows the RBFNN-GA algorithm architecture. The state space shrinks continuously during evolution and finally shrinks to a small attracting subset, and each attracting subset has a certain attraction domain. The energy function is a basic quantity of this type of network. Using the
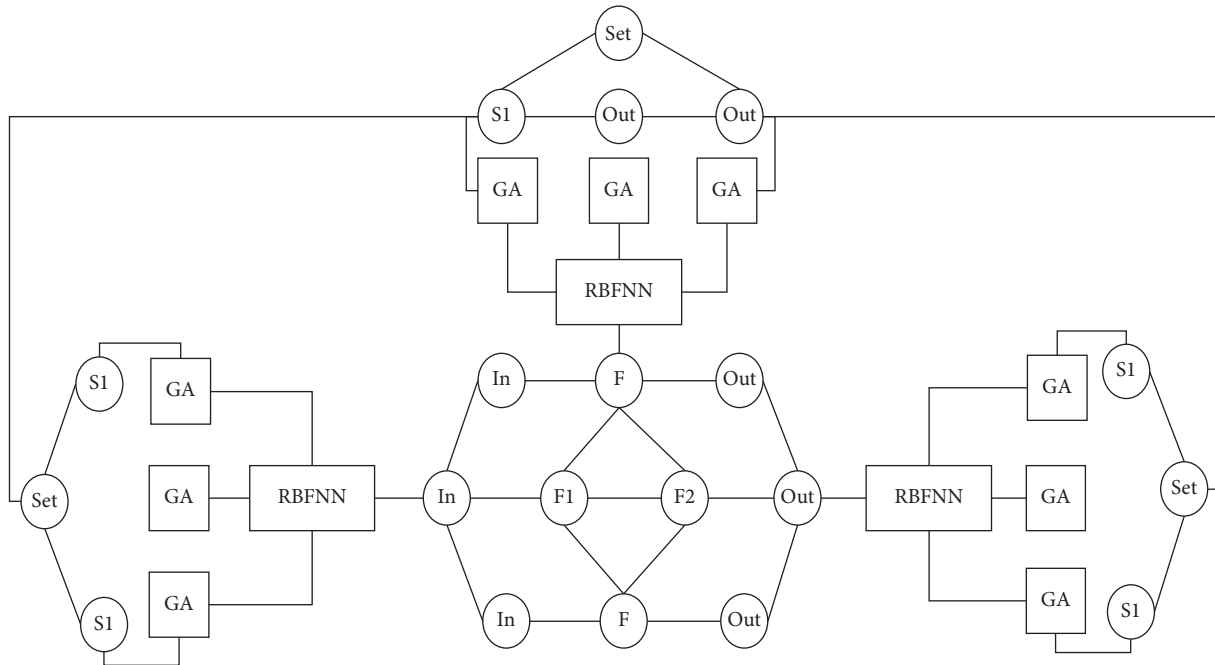
Figure 3: RBFNN-GA algorithm architecture.

local minima of the energy function, operations such as associative memory, information compression, and coding can be performed: only the global minima of the energy function can be used to solve combinatorial optimization problems, such as TP problems, visual matching, and problem-solving. Output layer learning is a kind of supervised learning. In supervised learning, the attributes of the pattern samples to be classified are known, and for each sample input and output, there is a corresponding guidance signal that matches its nature. Based on the network output supervision, the various objective function criteria of the signal and the actual output adjust the weights accordingly until the accuracy reaches the best requirements. The database is based on the implementation project, the standard map sheet is the spatial index, and the geological disaster professional data are classified as the first-level processing object. According to the storage form, it is divided into two parts: the graphic database and the attribute database. Respective graphic elements and associated fields of records (or unified numbers) realize the two-way dynamic connection, combined with multimedia database and graphical legend data, to lay the foundation for environmental geological survey and evaluation as the application goal. The specific operation is carried out in strict accordance with the input direction and input sequence in the technical requirements of spatial database construction and the method of distinguishing different line elements with different colors to form a comprehensive graphic line file in order to facilitate the hierarchical extraction of line files.

2.4. Delineation of Regional Features.

The Permian and Triassic strata are the most widely distributed in the area. Among them, the Triassic strata are the most well developed.

Carbonate rocks account for about 68% of the total thickness of the strata. Carbonatite and clastic rocks are multilayered in vertical distribution and are also called interbedded carbonate rocks. The interbedded carbonate rocks are distributed in a strip-like plane. That is to say, except for the Yanshan Movement, which is an obvious fold orogenic movement, the rest of the crustal movement is slowly ascending and descending. Therefore, the regional crust is relatively stable. In order to apply the projection transformation of GIS, we first transform the data format in the database. Figure 4 shows a quantitative histogram of the susceptibility of geological disasters.

The grid map of the susceptibility of geological hazards is obtained through spatial analysis and superposition based on the changing trend of geological hazard points. The stability of regional landslides, avalanches, ground subsidences, and unstable slopes is divided into three categories: unstable, basically stable, and stable using expert assessment methods; for debris flows, the integral method is used, and the evaluation factor is determined to be 15 items. The integral value method is used to assess the susceptibility of debris flow geological hazards. Taking the 1 : 100,000 topographic map of the study area as the basic base map, it is divided by the raster data processing method. Taking into account the geological environment and disaster development status of an area, the standard area grid is 2 km × 2 km. The area is divided into 66 rows and 6l columns, with a total of 1,856 cells. Compared with traditional computer remote sensing classification, its classification accuracy is significantly improved, providing technical reserves and references for land use dynamic monitoring and land use management. Using the face-to-point intersection function of GIS spatial analysis, the face-to-point intersection calculation is performed between the surface file of the divided grid and the point file of various disaster points (with the attribute of
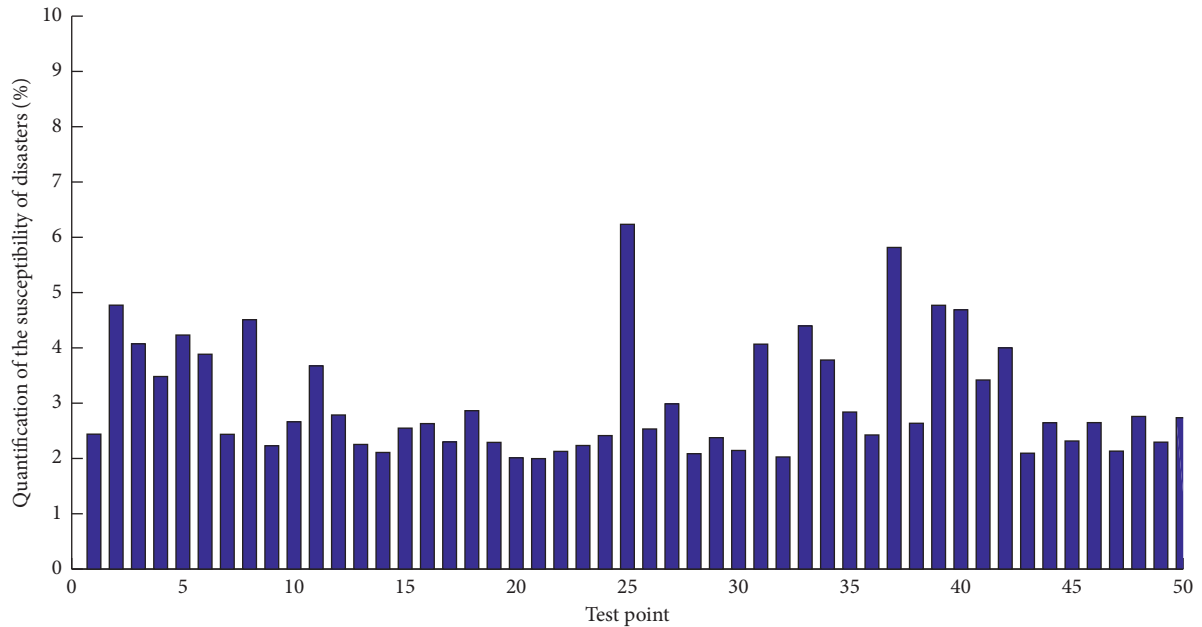
FIGURE 4: Quantitative histogram of the susceptibility of geological disasters.

proneness degree) to obtain different levels of disasters. Each cell is assigned a value based on the highest level of a type of disaster point that appears in the range. Using the method of GIS spatial analysis, the surface files of the divided grid and the point files of various geological disaster points (with the attribute of proneness degree) are calculated to obtain the partition files of different levels of various disasters. The partition files of different levels of various disasters obtained above are superimposed according to the following formula, which is the result of the scoring area. First, we superimpose the files in various unstable or high-prone areas. When there are two or more geological hazard high-prone areas overlapping, the value is 5. Second, we intersect all kinds of basically stable or medium-prone surface files with the merged unstable or high-prone surface files, subtract the common parts, and superimpose and merge according to the above formula to obtain the medium-prone layer. Third, we intersect all kinds of stable or low-prone surface files with the merged high- and medium-prone files and subtract the intersecting part, superimpose, and merge according to the above formula to get the low-prone layer. Finally, we subtract the high-, medium-, and low-prone areas from the whole picture, and the remaining areas are less prone areas. After merging, the zoning grid map of the susceptibility of the whole area can be obtained.

3. Application and Analysis of GIS Technology Based on RBFNN-GA Algorithm in the Delineation Model of Geological Hazard Prone Areas

3.1. RBFNN-GA Algorithm Data Extraction. The hardware requirements of this system are as follows: CPU Pentium 1.5 GHz, memory 512 M, hard disk 20 G and above

running on computers, Windows 2000/XP/2003 operating system, and software platforms that mainly include ERDAS IMAGINE 8.7, MATLAB 6.5, and Visual Studio 2005. Through field surveys and referring to the current land use map, we select typical land use plots and record their latitude and longitude coordinates. We use the remote sensing image file and text file conversion tool provided by ERDAS IMAGINE to convert the established training sample area into a text file; use the text file reading function provided by MATLAB to write the band values of each land use type into the matrix; and build a network model through the newrb function. The results of the network simulation are output in the form of a matrix, and when the accuracy analysis and classification results are evaluated, the image map is used as the basic data. Therefore, the classification results in the matrix form should be converted and restored into remote sensing images. This work can be done through remote sensing.

Figure 5 shows the normalized comparison line chart of RBFNN network evaluation indicators. We integrate the normalized results of the aforementioned various evaluation index data analysis and then use all the indicator inputs by the spatial overlay and analysis function in ArcGIS to obtain the digital matrix result; in ArcGIS, the Raster Calculator distributes the index data in the study area according to the weight. After the stacking calculation, the results need to be reclassified. Using the natural discontinuity grading method in the classical tool in ArcGIS, the stacking results are divided into three levels, and the regional geological hazard risk assessment result map is obtained. The system function mainly includes the selection of training sample area and the establishment of RBFNN algorithm and simulation. According to the scores in the result map superimposed by various indicators of the region, the geological hazard risk is
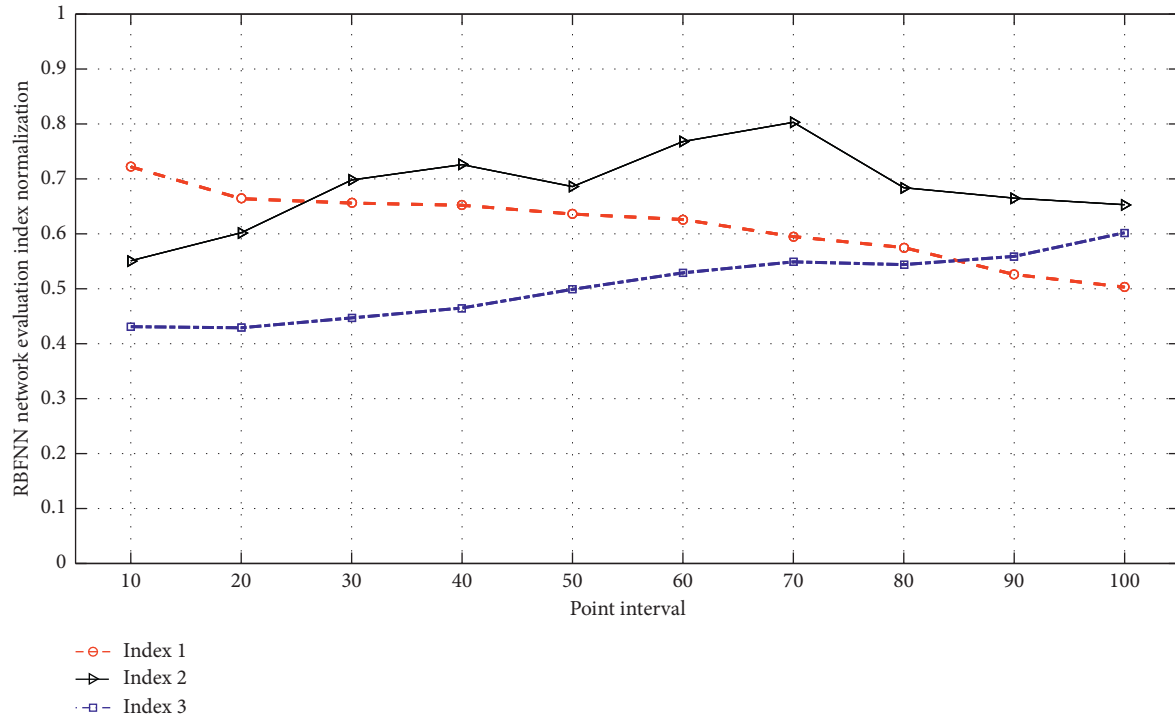
FIGURE 5: RBFNN network evaluation index normalized comparison line chart.

divided into three levels, namely, high, medium, and low, and then according to the actual situation, the repeated operation of grid or merging small redundant areas is performed. Finally, an artificially modified and integrated geological hazard risk zoning map can be obtained. The data calculated by GIS technology have been obtained by statistics of 16 villages and towns in the region. The requirement of consistency requires the establishment of a new judgment matrix through the relationships at all levels until the calculation and judgment results are reasonable. According to the principle of checking the consistency of the judgment matrix, matrix 1, matrix 2, matrix 3, and matrix 4 are calculated and tested, respectively, and the average CR < 0.1 is obtained. It is known that the constructed regional geological disaster risk assessment judgment matrix meets the consistency requirements.

3.2. Realization of GIS Delineation in Areas Prone to Geological Hazards. In this study, the neural network was set with 6 input nodes and 1 output node, and the initial value of the hidden node was 20. We used the control force to train the network parameters. The number of SGA implicit function nodes is fixed as 6, and the crossover probability is 0.7. The probability of variation is 0.05. The maximum fitness value changes during the training process as described in the text. As the core part of the geographic information system, the design quality of the database will not only affect the cost and quality of the system construction but also affect the operation, maintenance, and data update of the system. At the same time, the content and structure of the database determine the quality of the

system and will have a direct impact on the integration of GIS and other application technologies. When designing the system, we fully consider the scalability and compatibility of the system. In terms of the encoding of disaster-type information, the selection of the base map coordinate system, the design of the database, and the system interface functions, as much as possible, there is room for further development to facilitate the expansion of the system. With database migration, when new modules are added, the functional structure of existing modules and the entire system will not be greatly affected.

Figure 6 shows the distribution of curvature of geological hazards in the GIS system. According to the survey of geological disasters, the probability of geological disasters on concave slopes is relatively small, while the probability of geological disasters on straight and convex slopes is greater. The more curved the slopes of convex slopes are, the more concentrated the stress on the slope is. Therefore, for the slope curvature, when the curvature of a certain area is less than 0, the possibility of geological disasters is smaller. A large number of human engineering activities have a direct or indirect impact on the formation and development of landslides, collapses, and unstable slopes in the study area. Considering that traffic construction such as roads and railways is the most representative human engineering activity in the region, it has the most impact on geological disasters. Obviously, it has the characteristics of penetration and full coverage, so the analysis of human engineering activities in this study takes the regional secondary road as the baseline and makes buffer analysis. A value close to 0 means that the prior probability is very close to the conditional probability, and the certainty of the occurrence of
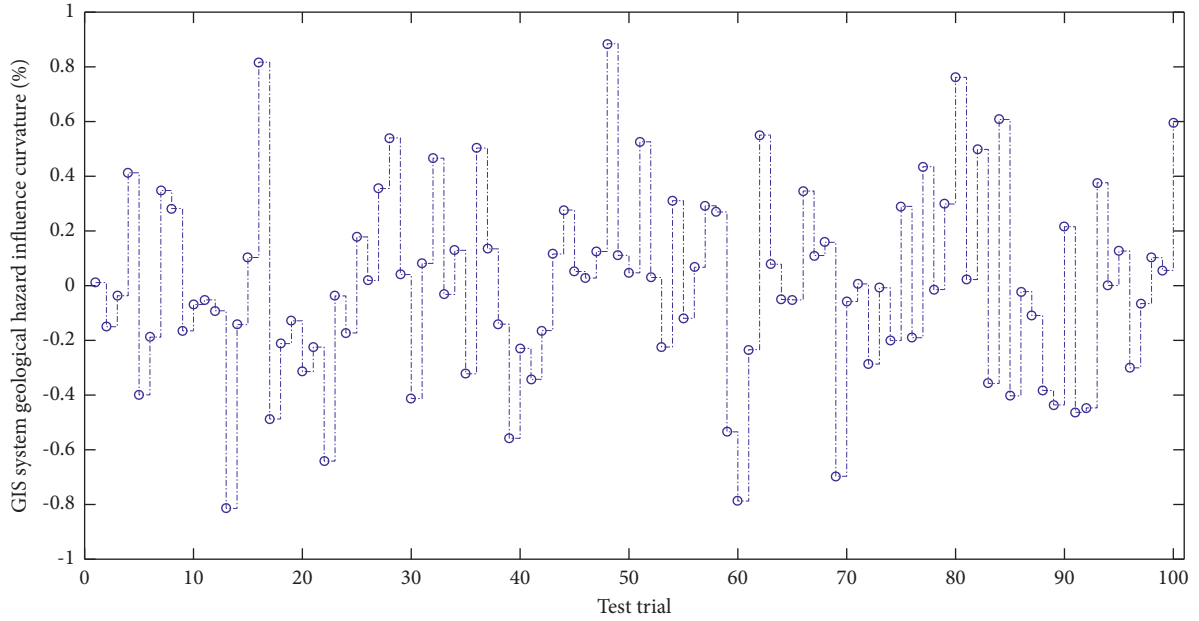
Figure 6: TheGIS system geological hazard influence curvature distribution.

the event is close to the regional average. When the curvature is greater than 0, the possibility of geological disasters is high, and the greater the curvature is, the more unstable the slope is. We use the Raster Calculator in ArcGIS to do the difference to calculate the slope height index map of the study area. Finally, it can be normalized by Raster Calculator in ArcGIS. Table 1 shows normalized processing of GIS disaster indicators.

The improved HANHGA-trained neural network controller and the SGA-trained neural network controller were used to control the double-stage inverted pendulum, and the simulation block diagram of the entire control system was built in Simulink using MATLAB. It can be seen that the improved HANHGA has a faster convergence speed and higher accuracy. The improved HANHGA training neural network has an error of 4.2% when it is trained for 100 generations, while the error of the SGA training neural network is only at the end of 200 generations. It reaches 1.2%, so the improved genetic algorithm has greater advantages than a simple genetic algorithm in convergence speed and in finding the global optimal solution. As a comparison between the improved HANHGA's RBFNN control method and the SGA neural network control method, it can be seen that the improved RBFNN control method can restore the inverted pendulum to a stable state faster, which proves the superiority of the improved algorithm. The data in the figure represent the control effect of the improved algorithm. It can be seen that it has a smaller overshoot and a shorter settling time, which proves the superiority of the improved algorithm.

3.3. Example Application and Analysis. On the ERDAS IMAGINE platform, using the Image Interpreter function provided by it, in the Utilities menu, the layer overlay module completes the overlay of each monochrome band,

Table 1: Normalized processing of GIS disaster indicators

Indicator index	Raster calculator	Weight	Curvature (%)
1	Landslides	0.32	13.5
2	Collapses	0.17	22.1
3	Slopes	0.42	31.2
4	Mudslides	0.09	26.4
5	Cracks	0.11	19.6

the data type is output as unsigned 8-bit data, and image fusion is mainly performed in the Spatial Enhancement menu. The fusion method uses a principal component (principal component analysis), and the resampling aspect is cubic convolution sampling. The key technologies for image geometric correction mainly include the selection of digital geometric correction calculation models and ground control points. For the determination of the gray value of the digital image and the resampling of the gray value of the digital image, the selection of the resampling method is particularly important. The commonly used gray value resampling methods mainly include the nearest point method, the bilinear interpolation method, and the cubic convolution method. The essence of the nearest point method resampling is to take the gray value of the known pixel point closest to the conjugate position in the original distorted image as the gray value of the output pixel. Figure 7 shows the gray value interpolation of the output pixel fitting.

Based on the powerful spatial analysis capabilities of GIS, for various geological disaster risk evaluation indicators in a certain area, it is necessary to normalize each indicator to achieve the final comprehensive risk evaluation. In this study, DEM images are downloaded from the geospatial data cloud of the National Academy of Sciences. We use the ArcGIS automatic stitching function to get the DEM map. After SPSS software completes logistic regression analysis,
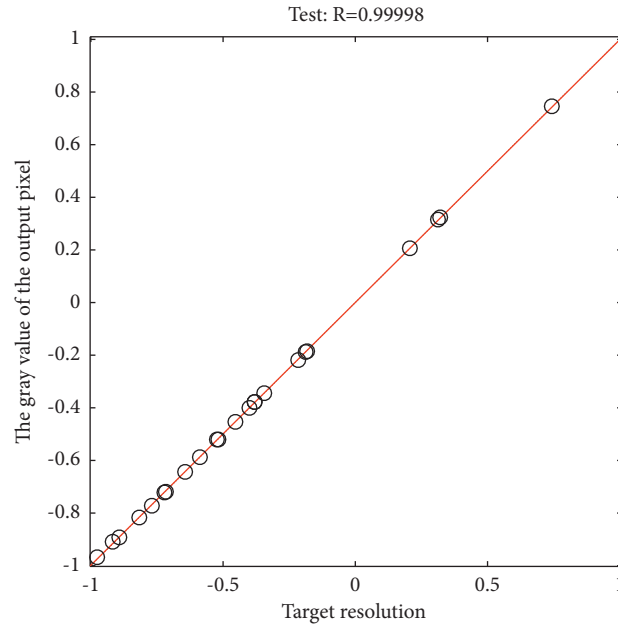
FIGURE 7: Gray value interpolation of output pixel fitting.
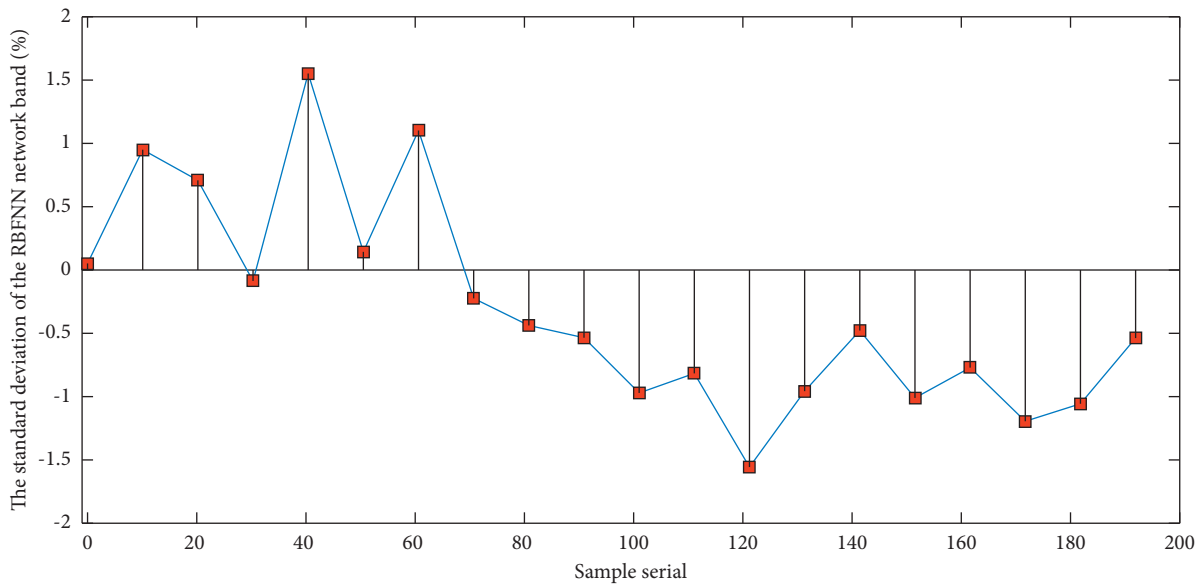
FIGURE 8: The standard deviation line chart of the RBFNN network band.

the output digital matrix can also be converted into raster graphics in GIS software, thus solving the problem of mutual import of data in GIS and professional statistical analysis software. In order to unify the coordinates and facilitate the postprocessing, the DEM is converted into general coordinates. According to the distribution of geological hazards, we use Spatial Analyst in ArcGIS density tool to designate a circular area and use a field radius of 1,000 m to analyze the point density of geological hazards. The higher the point density in the area is, the more it reflects, the higher the probability of geological disasters in this area, and the greater the risk of geological disasters is. The data sample set is divided into two parts: the modeling sample set (1,056 data

samples) used to build the RBF neural network model and the test sample set (168 data samples). Through calculation, the performance value convergence curve can be drawn when constructing the network. It can be seen that it can converge to the target value after 126 generations.

Figure 8 shows the standard deviation line chart of the RBFNN network band. The data show that the order of the standard deviation of each band is shown in the figure. It is generally believed that there is good agreement between the amount of information of remote sensing images and the image standard deviation; that is, the larger the standard deviation is, the richer the information is. The standard deviation of each band of the image is greater than 3, the

largest is 18, the smallest is 3.4, and the average is 9.4, indicating that the information contained in the features is more enriched, and it is feasible to extract land types. The social attribute factors mainly include important engineering facilities such as reservoirs and a small number of oil extraction plants. In some areas, there are a small amount of urban construction and roads, and human engineering activities are relatively weak, so its social attributes are low. The development density of geological disasters and landslides in this area is about 0.3 places/$100\,km^2$. Among them, there are only 11 unstable slopes in the southern part of the area, which are of low risk. Considering various factors, the geological hazard risk in this area is positioned as medium, which is the risk area in the regional risk assessment.

4. Conclusions

This study comprehensively analyzes the main influencing factors that induce regional geological disasters. Based on the theoretical experience and field investigations, the evaluation index system for geological hazards and the regional geological environment quality evaluation index system are established, and the quality characteristics of the regional geological environment are determined. According to the characteristics of the regional geological disaster risk assessment problem, two simple and practical mathematical methods, GIS technology and neural network, are selected to establish the RBFNN-GA algorithm evaluation model, and the final evaluation results are analyzed and compared. Based on ArcGIS as the underlying platform, a spatial graphic database is established. The attribute database is linked by an external database. The evaluation is obtained through secondary development on the basis of the geological spatial database. The analysis module realizes the spatial analysis function. Finally, the fuzzy comprehensive judgment method is used to optimize and select the evaluation index, and the neural network method is used to determine the weight of the evaluation index. A set of scoring standards based on quantitative indicators is formed, and membership functions are used to obtain values; qualitative indicators are evaluated using expert scoring methods to quantify indicators. At the same time, the main geological disasters in the region are the evaluation objects, combined with GIS technology to comprehensively evaluate the susceptibility and zoning of regional urban geological disasters, and the study area is divided into four levels: excellent, good, medium, and poor. The result shows that the evaluation model is reasonable and the system has strong practicability.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] S. K. Chandar, "Hybrid models for intraday stock price forecasting based on artificial neural networks and meta-heuristic algorithms," *Pattern Recognition Letters*, vol. 147, pp. 124–133, 2021.

[2] Q. Tan, M. Bai, P. Zhou, J. Hu, and X. Qin, "Geological hazard risk assessment of line landslide based on remotely sensed data and GIS," *Measurement*, vol. 169, Article ID 108370, 2021.

[3] Q. Tan, Y. Huang, J. Hu, P. Zhou, and J. Hu, "Application of artificial neural network model based on GIS in geological hazard zoning," *Neural Computing and Applications*, vol. 33, no. 2, pp. 591–602, 2021.

[4] B. S. Sun, Z. Chen, B. Q. Gu, and X. D. Huang, "Injection molding process optimization of multi-objective based on MUD-RBFNN-GA," *Applied Mechanics and Materials*, vol. 37-38, pp. 564–569, 2010.

[5] S. Deng and W. Li, "Spatial case revision in case-based reasoning for risk assessment of geological disasters," *Geomatics, Natural Hazards and Risk*, vol. 11, no. 1, pp. 1052–1074, 2020.

[6] G. Chen, "GIS method and its application for harmonious evaluation of urban construction land and geological environment," *Arabian Journal of Geosciences*, vol. 12, no. 19, pp. 1–6, 2019.

[7] W. Zhigang, T. Jinhua, L. Guorong, and Z. Qun, "Application of GIS rapid mapping technology in disaster monitoring," *Remote Sensing and Spatial Information Sciences*, vol. 42, p. 3, 2018.

[8] X. M. Zhao, D. S. Huang, and Y. Cheung, "A novel hybrid GA/RBFNN technique for protein sequences classification," *Protein and Peptide Letters*, vol. 12, no. 4, pp. 383–386, 2016.

[9] S. H. Mai, M. E. A. B. Seghier, P. L. Nguyen, and J. Jafariasl, "A hybrid model for predicting the axial compression capacity of square concrete-filled steel tubular columns," *Engineering with Computers*, pp. 12–18, 2020.

[10] H. He, D. Hu, Q. Sun, L. Zhu, and Y. Liu, "A landslide susceptibility assessment method based on GIS technology and an AHP-weighted information content method: a case study of southern Anhui, China," *ISPRS International Journal of Geo-Information*, vol. 8, no. 6, p. 266, 2019.

[11] B. Liu, X. Chen, Z. Zhou, M. Tang, and S. Li, "Research on disaster resilience of earthquake-stricken areas in Longmenshan fault zone based on GIS," *Environmental Hazards*, vol. 19, no. 1, pp. 50–69, 2020.

[12] K. Peng, M. N. Amar, H. Ouaer, M. R. Motahari, and M. Hasanipanah, "Automated design of a new integrated intelligent computing paradigm for constructing a constitutive model applicable to predicting rock fractures," *Engineering with Computers*, pp. 11-12, 2020.

[13] H. A. Afan, M. F. Allawi, A. El-Shafie et al., "Input attributes optimization using the feasibility of genetic nature inspired algorithm: application of river flow forecasting," *Scientific Reports*, vol. 10, no. 1, pp. 4684–4715, 2020.

[14] B. Yuxin, "GIS and deep learning for geology and geomorphology," *Advances in Intelligent Systems and Computing*, vol. 1, pp. 199–204, 2021.

[15] L. Qin, S. Feng, and H. Zhu, "Research on the technological architectural design of geological hazard monitoring and

rescue-after-disaster system based on cloud computing and Internet of things," *International Journal of System Assurance Engineering and Management*, vol. 9, no. 3, pp. 684–695, 2018.

[16] J. Tao, Z. Yu, R. Zhang, and F. Gao, "RBF neural network modeling approach using PCA based LM-GA optimization for coke furnace system," *Applied Soft Computing*, vol. 111, Article ID 107691, 2021.

[17] P. C. Sahu, S. Panda, S. P. Panigrahi, and K. Rarvathi, "RBFNN equalizer using shuffled frog-leaping algorithm," *Soft Computing: Theories and Applications*, pp. 549–558, Springer, Singapore, 2018.

[18] N. Ahmadi, M. Nilashi, S. Samad, T. A. Rashid, and H. Ahmadi, "An intelligent method for iris recognition using supervised machine learning techniques," *Optics and Laser Technology*, vol. 120, Article ID 105701, 2019.

[19] D. Li, Y. Gong, G. Tang, and Q. Huang, "Research and design of mineral resource management system based on big data and GIS technology," in *Proceedings of the 2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, Xiamen, China, May 2020.

[20] S. Zhang, Y. Wang, and T. Chang, "Assessment of geological hazards in ningde based on hybrid intelligent algorithm," *Sensors and Materials*, vol. 30, no. 3, pp. 565–575, 2018.

[21] A. Anarghya, N. Rao, N. Nayak, A. R. Tirpude, D. N. Harshith, and B. R. Samarth, "Optimized ANN-GA and experimental analysis of the performance and combustion characteristics of HCCI engine," *Applied Thermal Engineering*, vol. 132, pp. 841–868, 2018.

[22] T. Liu, "Natural suitability evaluation of habitat environment in transition zone based on GIS spatial analysis technology—taking Beichuan county as an example," *Journal of Geoscience and Environment Protection*, vol. 9, no. 8, pp. 151–162, 2021.

[23] Z. Wang, Q. Liu, and Y. Liu, "Mapping landslide susceptibility using machine learning algorithms and GIS: a case study in Shexian county, Anhui province, China," *Symmetry*, vol. 12, no. 12, p. 1954, 2020.

[24] Z. Liu, Y. Wang, X. Hua, H. Zhu, and Z. Zhu, "Optimization of wind turbine TMD under real wind distribution countering wake effects using GPU acceleration and machine learning technologies," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 208, Article ID 104436, 2021.

[25] G. Chen, J. Zhao, L. Yuan, Z.-j. Ke, M. Gu, and T. Wang, "Implementation of a geological disaster monitoring and early warning system based on multi-source spatial data: a case study of Deqin county, Yunnan province," *Natural Hazards and Earth System Sciences Discussions*, vol. 2017, no. 7, pp. 12–15, 2017.

[26] Y. Xia, Y. Wang, S. Du, X. Liu, and H. Zhou, "Integration of D-InSAR and GIS technology for identifying illegal underground mining in Yangquan district, Shanxi province, China," *Environmental Earth Sciences*, vol. 77, no. 8, pp. 15–19, 2018.

[27] K.-k. Shang, T.-c. Li, M. Small, D. Burton, and Y. Wang, "Link prediction for tree-like networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 6, Article ID 061103, 2019.

Research Article

An Empirical Analysis of Corporate Financial Management Risk Prediction Based on Associative Memory Neural Network

Hui Chu (ID)

China University of Mining and Technology-Beijing, School of Management, Beijing 100083, China

Correspondence should be addressed to Hui Chu; bqt2000501002@student.cumtb.edu.cn

As a human brain-like computational model that can reflect the cognitive function of the brain, the problem of dynamic analysis of associative memory neural networks has attracted the attention of scholars. This paper combines associative memory neural networks with enterprise financial management risks, studies the synchronization control and stability analysis problems of unidirectional associative memory-like human brain amnestic neural networks with perturbation and mixed time-varying time lags, proposes a bidirectional associative memory-like brain stochastic amnestic neural network model with mixed time-varying time lags, designs a discrete-time sampling control strategy based on the model, and studies various types of recent financial risks. Based on the early warning research, based on the associative memory neural network method, we propose to reconstruct the risk categories, including improving the enterprise risk management system, enhancing the awareness of financial risk management from top to bottom, and strengthening the core competitiveness of the enterprise itself and control measures for financing, investment, operation, and cash flow risks.

1. Introduction

Associative memory neural network is by imitating the working pattern of neuronal cells in the human brain; firstly, the pattern to be memorized is stored in the form of weight network of the neural network; when receiving the information of incomplete or defective pattern from outside, the neural network at this time makes the input pattern continuously change its value and converge to the memorized pattern by massive parallel computation. The neural network has a good robust performance, which means that the associative memory neural network has a good fault tolerance performance. The patterns to be memorized are first stored in the form of a network of weights of the neural network. When receiving information about incomplete or fragmented patterns from outside, the neural network then continuously changes its values and converges to the memorized patterns through massive parallel computation. Associative memory neural networks are now widely used in pattern recognition, image processing, etc. Self-associative memory means that the input aberration patterns are equal

to the remembered patterns; in contrast to heteroassociative memory, the input aberration patterns are not the same as the remembered patterns but show corresponding mapping relationships. Self-associative memory refers to the application of associative memory in which the initial pattern of input is identical to the remembered pattern of output. The first step in self-associative memory is to deposit the desired memorized pattern (e.g., the word computer) in a specific form into the network weights through the learning algorithm of a neural network [1]. Then, it is possible to input the pattern with cue information, and the neural network performs continuous iterative operations to give the correct output, and this cue information is not the complete word but has noisy information, but the neural network is still able to calculate to get the remembered word information; i.e., this associative memory neural network has a certain degree of fault tolerance. In fact, fault tolerance is one of the criteria to judge whether the application of associative memory is feasible. The memory process of the neural network, that is, the setting of the weights of the neural network, needs to translate the memorized patterns into the form of the state

values of each neuron when the neural network is in a stable state; that is, generally, a stable point of the neural network can store a pattern, and the determination of the value of the stable point needs to be set according to the necessary conditions required by the associative memory neural network model adopted in this chapter, generally by substituting the memory pattern into the set of differential equations of the neural network, then listing the corresponding inequality based on the above necessary conditions, and relating this inequality to the system of differential equations coupled to solve a specific set of solutions and eventually determine a more suitable weight [2]. This process is also performed using specific training algorithms (e.g., through network self-feedback). Since the self-associative memory input and output patterns are identical, meaning that both have the same dimensionality, this is a characteristic of self-associative memory.

Under the conditions of market economy, enterprises must be aware of the objectivity of financial risks, and at present, enterprises generally have a weak awareness of financial risks, blind fund-raising, imprudent investment, inadequate financial risk management system, and state-owned enterprises, which have weak points such as imperfect financial risk evaluation mechanism and lack of risk control measures. Thus, studying the characteristics of enterprise financial risks and taking certain measures to prevent and manage them has become an important issue facing us now. Although the rapid development of various industries also brings many risks, the industry frequently appears in the financial crisis and even eventually leads to bankruptcy of enterprises [3]. The determination of the stability point value is set according to the necessary conditions required by the associative memory neural network model adopted in this chapter, generally by substituting the memory model into the set of differential equations of the neural network and then listing the corresponding inequalities according to the above necessary conditions. This inequality is then combined with the set of differential equations to solve a series of specific solutions and finally determine a more suitable weight value. Therefore, the financial risk management of enterprises is very important and can determine whether the enterprise can develop in a high speed and sound manner. In recent years, scholars have also increased their research on the financial risk management of enterprises, but a systematic theoretical system has not yet been formed for the study of the financial risk of enterprises. Furthermore, the focus of existing research is to analyze the financial risks of enterprises and put forward financial risk control opinions, and there are relatively few studies that analyze the steps of financial risk management of enterprises, i.e., the identification, evaluation, and control means of financial risks, in-depth. Based on associative memory neural network, this paper combines the relevant theoretical literature on enterprise financial risk control with the specific case of financial risk control of the Storm Group and researches the identification, analysis, and evaluation of financial risk and control of enterprises, hoping to play a supplementary verification role to the current theoretical research on financial risk control in enterprises.

2. Status of Research

Artificial neural networks have been intensively studied in recent decades due to their wide application in many fields such as pattern recognition, associative memory, signal processing, and optimization. Some applications, such as those in optimization theory, require neural networks to have only a single stable point. However, some other applications such as associative memory and pattern recognition require the existence of multiple stable equilibria in the neural network itself.

The multistability of two recurrent neural networks with phase plane origin symmetric activation functions is studied in the literature [4]. Multiple μ-stability of neural networks with unbounded time-varying time lags is considered in the literature [5]. The multistability of recurrent neural networks with nonmonotonic activation functions and unbounded time-varying time lags is discussed in the literature [6]. In the literature [7], the multistability of fractional-order neural networks in the Mittag-Leffler sense with segmental constant parameters is discussed. The multistability of a class of Hopfield neural networks with stochastic time lags was studied in the literature [8] by applying Schauder's immobility point principle and the related stochastic time-lag theory, where there exist $2n$ rectangular invariant set regions under certain conditions and at least one of these regions has an equilibrium point, and by proposing two sufficient conditions to ensure that these equilibrium points are stable. A new condition that is less restrictive relative to the Lipschitz continuum condition in complex-valued activation functions is proposed in the literature [9], and a series of criteria are presented to guarantee the existence, uniqueness, and numerical examples of global asymptotic stabilization points of complex-valued recurrent neural networks. In the literature [10], the stability of recurrent neural networks with time-varying time lags is analyzed using FTM (Flexible Edge Method), and several new stability criteria are proposed to describe the stability of this neural network by constructing a new type of Lyapunov function. In the literature [11], a novel design idea for CVHAMs (complex-valued multistate Hopfield associative memory) is proposed to analyze the stability of CVHAM systems utilizing the energy function method and finally obtain that the network can converge to an immobile point for an arbitrary input value, and the projection geometry of the GPR (generalized projection rule) is discussed. According to the literature [12], Internet enterprises have the financial characteristics of the "flow economy" model, asset-light operation model, cost-light and expense-light, fast product replacement, and equity financing. The financial risk of Internet enterprises has both the characteristics of general enterprises and its special characteristics. Compared to the general enterprises, the risk of capital shortage of Internet enterprises is the most significant. The literature [13] introduced the EVA valuation model to explore the link between financial risk and an enterprise value of Internet enterprises, used the EVA model and the traditional valuation model to assess the company value of the case company NetEase, and

concluded that the EVA valuation model is more accurate. The literature [14] argues that Internet enterprises mainly have risks in profitability model, cash flow, financing, investment, and government regulation. It also suggested that Internet enterprises should establish a sound financial risk early warning mechanism, set up a professional auditing institution and audit system, strengthen comprehensive budget work, improve the accounts receivable control mechanism, and strengthen financing and investment risk management. The literature [15] argues that the assurance of asset profitability and the control measures proposed to prevent unexpected losses are accomplished through risk management. Literature [16] argues that companies should first establish a sound internal control system when making business decisions, which can, to a certain extent, avoid suffering financial risks and causing operational difficulties when financing activities. Literature [17] studies the origin of the risk, the financial risk mechanism of action as the goal to explore, and more targeted risk identification, evaluation, and control, eliminating the hidden financial risk factors.

3. A Risk Model for Corporate Financial Management Based on Associative Memory Neural Networks

3.1. Associative Memory Neural Network Model. The basic steps of the current recurrent neural network training algorithm are to first preprocess the samples to be recognized to obtain numerical feature vectors, input these vectors sequentially into the network, then compare the output of the network with the going output by a difference, and use this difference as negative feedback to correct the network weights. The advantage of this algorithm is that it is simple to operate and the training process is unsupervised learning, while the disadvantage is also evident in that it needs to be retrained once every time a new sample pattern is remembered, which may consume more time in the case of larger network dimensions. The neural network model is built by adjusting the weights of its neurons after it receives data, and the input variables of the model can change according to the actual needs, such as market factors and financial indicators. The neural network model can be automatically trained to filter the best variables and adjust the weights according to their implied relationships to build a nonlinear model and improve the model accuracy. The training method is used to correct the network weights by continuously feeding in training samples so that the sample patterns to be remembered are embedded in the network weights; i.e., the network "remembers" the samples. It is possible to use algebraic methods to solve the weights directly, because the network has a large number of attractors and can remember a large number of samples, by using the samples as input, assuming that the output is the desired output, representing both as vectors, and substituting them into the differential equations of the system, using the knowledge of linear algebra to solve a set of weights network, then using other samples for input, and finally solving the

corresponding number of analytical solutions. The intersection of these analytic solutions is the network of weights to be determined if these solutions do not intersect. Then, it means that the memory capacity of the network is not sufficient and a new model needs to be redesigned to ensure that it has sufficient memory capacity [18]. Next, an example of the design of associative memory will be given, as shown in Figure 1 for an example of the associative memory process.

Here is an example of the detailed steps of associative memory design; a recurrent neural network can be described by the following set of differential equations:

$$\frac{\partial}{\partial W_{ij}^l} J(w, b) = \frac{1}{b} \sum_{i=1}^m J(w, b; x^i, y^i) + \lambda W_{ij}^l. \tag{1}$$

The network is a more classical model of a memetic recurrent neural network. Several character samples are used for the training of this network until the network is given a network of weights that satisfy the conditions that make the network remember the characters to be remembered. For the sample image data to be recognized, the image first needs to be subjected to a series of preprocessing steps, which are graying, smoothing, denoising, and finally, transforming into an input vector. All samples of data are grayed out, smoothed, denoised, and then transformed into an array of vectors of the following form:

$$f_\phi(p, x) = d_p\left(x - \frac{\lambda_{\theta_1}}{d_p(x)}\right) + d_p\left(x - \frac{\lambda_{\theta_2}}{d_p(x)}\right). \tag{2}$$

And the output vector of the network can be represented as an array of vectors of the form

$$(\alpha, \beta, \gamma) = \frac{1}{r} \sum_{i=1}^k \left(\frac{\alpha_i + \beta_i + \gamma_i}{3}\right). \tag{3}$$

Substituting the above two sets of vectors into the system equation (1), we get

$$\nabla_{b'} J(w, b; x_i^*, y) = \delta^{l+1} \frac{\partial J(w, b)}{\partial b_{ij}^l} + \lambda \alpha_{ij}^l, \tag{4}$$

where $x_i^* x_i *$ is the stable equilibrium point of the system with the i-th set of samples as input, and this value is chosen according to conditions such as the form of the activation function. A sufficient condition for this equation to have a solution is that the number of stable equilibrium points of the neural network is greater than or equal to the number of samples to be memorized by the claim; that is, the number of attractors should be large enough to be able to remember a sufficient number of patterns. The equation can be solved using singular value decomposition, the solution of the equation is also the vector of connection weights of the network, and these weights imply information about the patterns to be remembered. The above form is generally a self-associative memory weight solving process that can be specifically applied in the fields of character recognition, face recognition, etc. Then examples will be given to illustrate the design and the weight solving process of recurrent neural
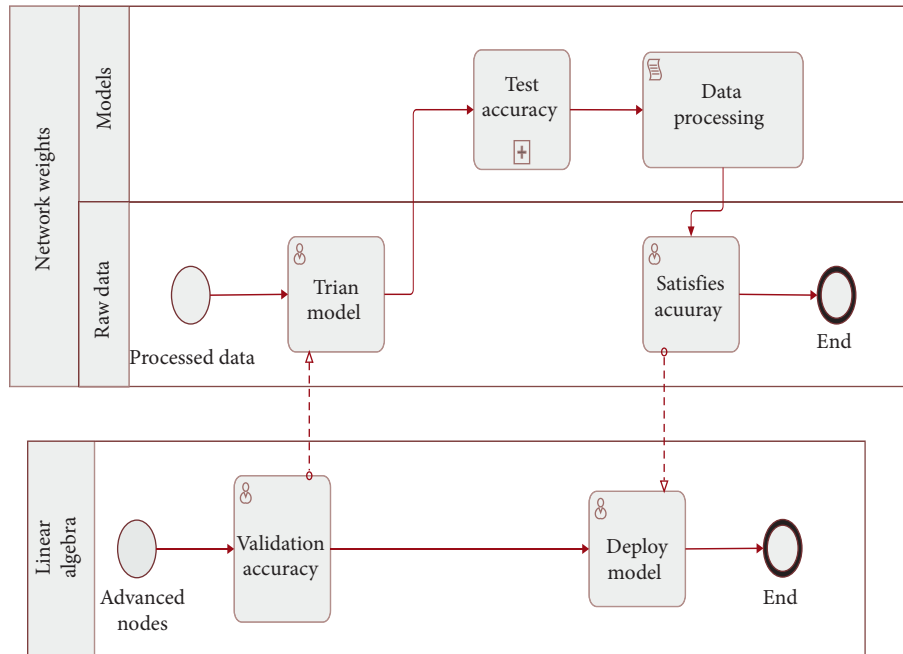
Figure 1: Example of the associative memory process.

network-based heteroassociative memory applications. Traditional character recognition algorithms include template matching and OCR [19]. For unbalanced sample data, you can start with downsampling and upsampling methods. The downsampling method, also known as the random downsampling method, refers to the random removal of the class of data with the most categories by the sampling method. The advantage is that the sampling method can improve the model accuracy when the deleted samples contain noisy data, and the disadvantage is that some important samples may be deleted. Oversampling method refers to synthesizing a part of small sample data by an algorithm. Compared with these character recognition algorithms, the recurrent neural network-based associative memory mentioned above is characterized by low design cost, simple processing logic, fast network convergence, less dependence on external conditions, and easy practical implementation. The most important point is that this associative memory system has good fault tolerance performance; high fault tolerance means that when there is more noise or disturbance in the sample, it can also correctly identify the target pattern, so if this associative memory method is applied to image recognition, although the current application is simple and the idea is more basic, after a little development, I believe there will be a use for it. The design steps of recurrent neural networks for associative memory are summarized as follows.

(1) Determine the model characteristics of the neural network to be used

(2) Convert the patterns to be remembered into vectors as inputs to the system as well as outputs to be substituted into the system equations

(3) The solution to the system of equations is obtained by appropriate choice of coordinates of the equilibrium point and by singular value decomposition

(4) Store the solutions of the system of equations as the weights of the neural network

(5) Build an associative memory neural network system and perform recognition tests.

The associative memory-based reconfigurable amnestic network circuit consists of four parts: the associative memory-based amnestic network circuit, the PRMC with input binary signals, the synaptic circuit block, and the control circuit. In the associative memory-based amnestic network circuit, the synaptic weights indicate the strength of the synaptic connections and thus only positive weights. The PRMC of the input binary signal can only be trained by the algorithm to achieve the corresponding function, so the synaptic circuit needs to represent negative weights, zero weights, and positive weights. In the associative memory-based reconfigurable memetic neural network circuit, the PRMC with input binary signals and the associative memory-based memetic neural network circuit uses the same synaptic circuit, and the synaptic circuit in the associative memory-based memetic neural network circuit is restricted to vary within the positive weights by a control circuit [20]. The synaptic circuits in the synaptic circuit block are shared, and these synaptic circuits can be used to construct both PRMCs with input binary signals and the associative memory-based memetic neural network circuits. As can be seen in Figure 2, the two subnetworks operate independently, so they can work in parallel.

By simulating learning and forgetting in associative memory, the reconfigurable memristor neural network
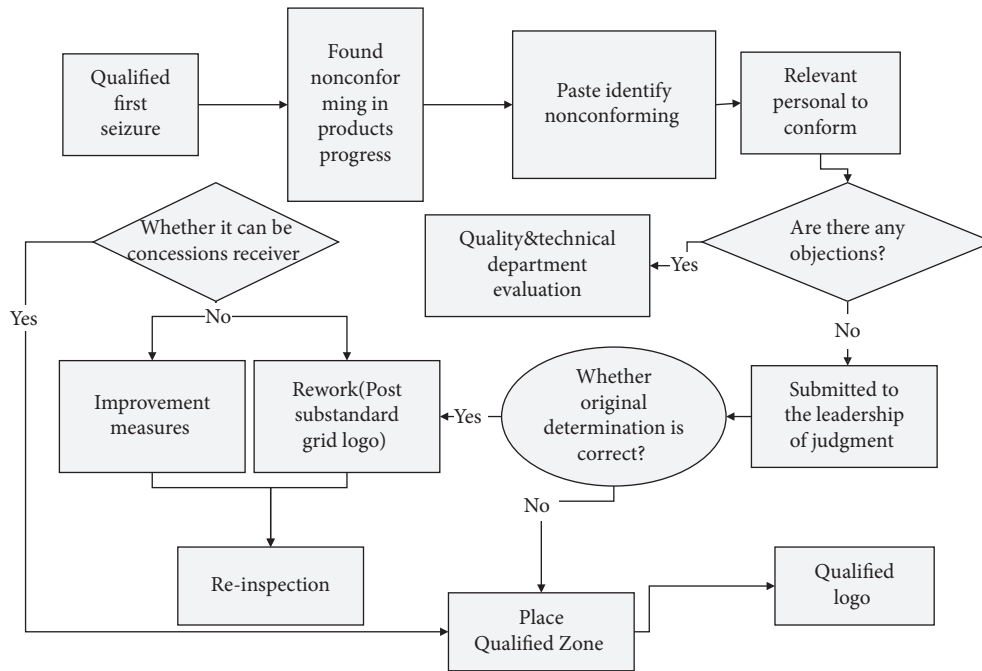
Figure 2: Structural diagram of reconfigurable amnestic network circuit based on associative memory.

circuit based on associative memory can dynamically change the circuit structure to achieve reconfiguration, corresponding to the following process: when the unconditioned stimulus signal and the conditioned stimulus signal are simultaneously input to the memristor neural network circuit based on associative memory, associative learning between the two is induced, so that the synaptic weight corresponding to the conditioned stimulus gradually increases. After the learning process is over, if the conditioned stimulus signal is kept input, the forgetting process is induced. During forgetting, the synaptic weight corresponding to the conditioned stimulus gradually decreases until it fails to activate the corresponding neuronal circuit. After the end of the forgetting process, the circuits of the associative memory-based amnestic network do not activate regardless of whether the conditioned stimulus signal is input [21]. And then, the synaptic circuits corresponding to the conditioned stimulus can be disconnected from the neuronal circuits, and these synaptic circuits can be fed into the synaptic circuit block. The synaptic circuits in the synaptic circuit block can be used to construct either PRMC or associative memory-based amnestic network circuits with input binary signals. In this way, circuit reconfiguration is achieved.

4. An Associative Memory Neural Network-Based Risk Prediction Model for Corporate Financial Management

Financial risk is a type of business risk that exists in financial activities such as raising, investing, spending, recovering, and allocating activities of funds in the process of social reproduction. And the risk is the deviation of expectations from reality. There is a distinction between financial risk in a narrow and broad sense. Financial risk in the narrow sense refers to the financial risk arising from the debt operation in the process of raising funds for the enterprise. Financial risk in the broad sense refers to the uncertainty of obtaining the expected financial results in the process of conducting the financial activities of the enterprise. Therefore, the financial risk exists in every enterprise and in the business activities of the enterprise and has a significant impact on the profit and loss situation and business conditions of the enterprise. Financial risk management is the prevention, control, and management of risk in enterprise financial management and is also a part of comprehensive enterprise risk management [22]. Internal diagnosis is a way for the enterprise itself to find out various problems in the business process through self-analysis and finally to solve them in a targeted manner. External diagnosis means that the enterprise hires an external third-party organization to analyze the financial operation of the enterprise. As a new management science theory, it is a management theory mainly developed by scholars according to the previous experience in risk management and financial management for all kinds of risks in the operation of enterprises. Financial risk management includes risk identification, assessment, analysis of causes, and control of various financial activities of the enterprise. To ensure the normal operation and capital movement of the enterprise and to avoid negative impact on the economic interests of the enterprise, the management process of financial risk management includes risk identification, assessment, analysis of causes and control of various financial activities of the enterprise, and timely and effective prevention and control measures based on its early warning role.

The most commonly used methods for financial risk evaluation include hierarchical analysis, efficacy factor method, and factor analysis. The basic principle of factor analysis is to group indicators with strong relevance into one category and replace each category with a factor, thus replacing all the original indicators with a few factors.

In this paper, the analysis of corporate financial management risk prediction is carried out by combining the associative memory neural network model with factor analysis. Factor analysis is chosen for the following reasons: firstly, factor analysis can reduce the number of original variables, by extracting and naming the main factors instead of the vast majority of the original information; secondly, the sample can be ranked and compared. According to the score of each main factor and the comprehensive score, the sample can be ranked, which not only shows the ranking of individual factors and clarifies which factors have a greater impact on the financial situation of the enterprise but also analyzes the comprehensive ranking of the factors after weighting, which is beneficial to the enterprise to clarify its strengths and weaknesses [23].

The steps of the factor analysis method are as follows:

(1) KMO and Bartlett's test were performed to determine whether the original variables were suitable for factor analysis; KMO was used to describe the magnitude of the correlation coefficient between variables, and the larger the value, the more suitable for factor analysis; while the smaller the value of Bartlett's test, the more suitable for factor analysis.

(2) Constructing factor variables and extracting principal component factors in place of all initial variables, the mathematical model for factor analysis is

$$f\left(\{x_1, x_2, \ldots, x_n\}\right) = g\left(h\left(x_1\right) + h\left(x_2\right) + \cdots + h\left(x_n\right)\right). \tag{5}$$

(3) Interpretation of the principal component factors is named, generally using the orthogonal rotation method to obtain the maximum loading values of the principal factors and clarify the meaning of what each factor represents.

(4) The expressions of each principal factor are derived from the component score matrix, and then a comprehensive evaluation model is constructed according to the respective weights, and the samples are ranked and analyzed.

$$C_{ij} = \sum_{i=1}^{n} x_i^2 \sigma_i + \sum_{j=1}^{n} x_j^2 \sigma_j + \sum_{k=1}^{n} x_k^2 \sigma_k. \tag{6}$$

Associative memory neural network is built based on decision tree model using histogram algorithm which is easier to segregate the data. The difference from previous decision tree models is that the associative memory neural network is oriented vertically, i.e., generating the leaves of the decision tree, whereas other decision tree models generate the levels of the tree, so the associative memory neural

network algorithm runs faster and stores fewer data. Its main feature is to traverse the entire training set and to make the attributes discrete for floating-point continuous variables; these k discrete data are constructed into a histogram with a specific width of k. The number of discrete values converged within each histogram is calculated. Since there are many components of financial risk, various factors and financial indicators will interact with each other to form a complex relationship. In this paper, relevant indicators are extracted, factors are extracted by significance test, and factors suitable for model building are obtained by sphericity test, principal component extraction, and other steps. Based on the obtained factors, the logistic model of stepwise regression is used to model the financial early warning. The modeling process gradually analyzes the factors of risk composition, using the advantages of the model to provide a more intuitive basis for the future decision-making of managers. The selected early warning indicators should be comprehensive; i.e., indicators should be selected from the traditional aspects of solvency, profitability, development capability, and operating capability, and indicators should be selected from the aspects of research and development capability that reflect the characteristics of the industry, making the range of indicators selected for the communication equipment manufacturing industry more comprehensive. The company invests capital in the production of products and recovers the capital and profits through the sales of the products. Inventory turnover can only be increased if sales are successful and inventory is cleared quickly, so inventory turnover indicates how quickly money can be recovered from the sale of goods. In general, a company can improve its liquidity by increasing its inventory turnover. The inventory turnover ratio indicates the level of inventory, while the current asset turnover ratio reflects the speed of turnover of current assets, which is the most liquid of all the assets of a company. A lower level of current asset turnover can have a greater impact on a firm's short-term repayment ability. The higher the current asset turnover ratio, the lower the relative financial liquidity risk. A slow turnover rate will require supplementary liquidity to participate in the turnover, which will create a waste of funds and reduce the profitability of the enterprise. Based on the above theory, in the subsequent classification, the optimal cut-off point can be found only according to the width of the histogram. The idea of the histogram algorithm is mainly reflected in the conversion of floating-point data into binary data, and the specific operation is to determine the number of buckets contained in each feature, update the data of each bucket separately after equal division, and substitute the features of enterprise financial management risk prediction into the associative memory neural network model, which is represented graphically as shown in Figure 3.

Compared to other models built on decision tree algorithms, associative memory neural networks are faster mainly in terms of running speed, while consuming less memory, and accuracy is not compromised, perfectly combining both fish and bear's paw. With the desire to go to the next level, the model can be optimized in the following two ways: to speed up the running speed, the original data
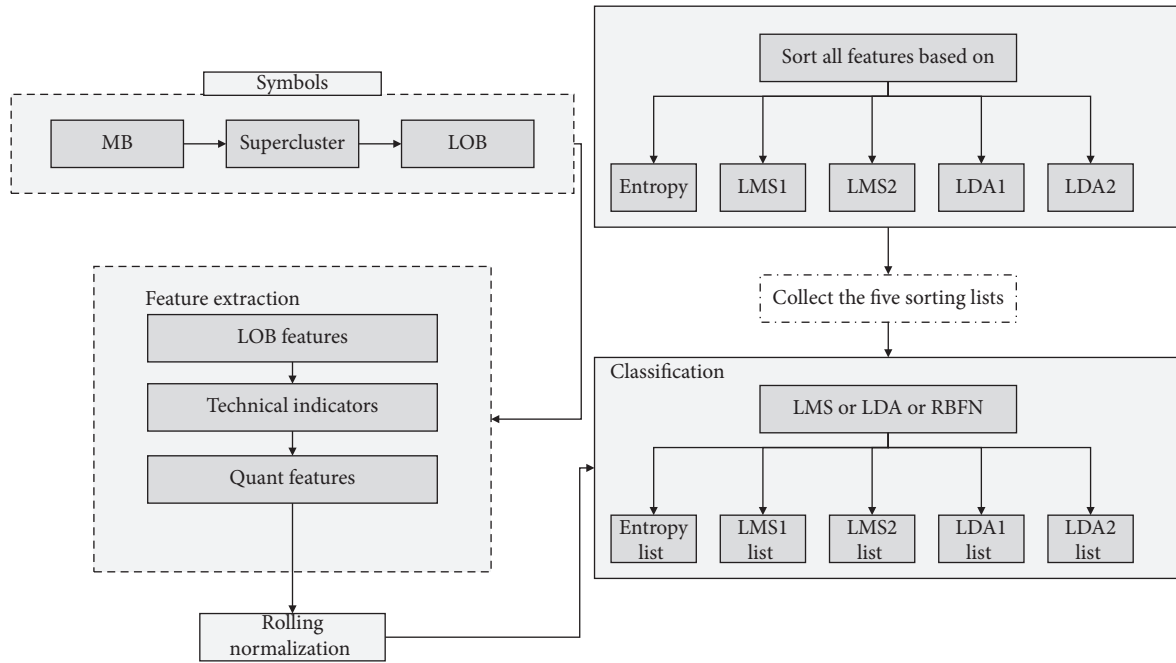
FIGURE 3: Flow chart of an interactive product styling model.

can be simply processed; reducing the number of features and data, converting feature variables into category features, or saving the data files as binary files, changing the training method of the model to parallel can also speed up the running speed of the model; to improve the accuracy of the model and reduce the model learning rate, starting from the model, mesh tuning the parameters of the model, choosing the best combination of parameters, increasing the learning times of the model, and making the model better understand the laws between the data can improve the accuracy of the model. Starting with the data, increasing the number of training data and preprocessing the data to solve the missing values and imbalance in the original data can also train the model better and improve the performance of the model.

4.1. Experimental Verification and Conclusions. The CART decision tree model with prepruning and postpruning design will be trained, and the decision tree graph and variable importance graph will be generated after the model training to better interpret the CART decision tree model discriminant results. Firstly, the discriminant results and ruleset of the overall financial risk assessment model of the enterprise are constructed. Because there are too many values of assessment features in the overall financial risk assessment of the enterprise, the variable importance graph in the overall financial risk assessment of the enterprise takes the top 5 importance variables, and the overall financial risk assessment model of the enterprise generates the variable importance graph as shown in Figure 4.

As can be seen from Figure 4, among the 18 characteristics of the CART decision tree model, "interest earned multiple," "total assets return," "accounts receivable turnover," "total assets turnover," and "weighted return on net

assets" ranked the top five in importance, with "interest earned multiple" being the most important indicator with the importance of over 0.5. Operating capacity reflects the efficiency and effectiveness of an enterprise's capital operation with different assets, and the turnover efficiency of different types of assets is usually used to determine the operating level of an enterprise. This indicates that enterprises should focus on "interest earned multiple" when assessing the overall financial risk, and the sum of the top five importance scores exceeds 90%; therefore, it also indicates that listed companies should focus on these five index characteristics in the process of assessing the overall financial risk of enterprises. The interpretation of the rule set generated by the decision tree model will provide quantitative evidence for the listed company to assess the overall financial risk of the company, which will help the listed company to measure whether it has overall financial risk.

The importance of the variables generated by the enterprise operational risk assessment model is shown in Figure 5. As can be seen from Figure 5, the importance of the six characteristics of the CART decision tree model is the same, with the importance of "operating profit margin" becoming the most important assessment indicator. Therefore, the model results indicate that listed companies should focus on these 6 indicators in the process of enterprise business risk assessment. Rule 1: when the "operating profit margin" is ≤−6.151, the enterprise has business risks. Rule 2: when "operating profit margin" >−6.151 and "accounts receivable turnover ratio" >3.203, the company does not have financial risk. Rule 3: when "operating profit margin" >−6.151, "accounts receivable turnover" ≤3.203, and "cost margin" ≤7.552, the enterprise, there is financial risk. Rule 4: when "operating profit margin" >−6.151, "accounts receivable turnover" ≤3.203, and "cost margin"
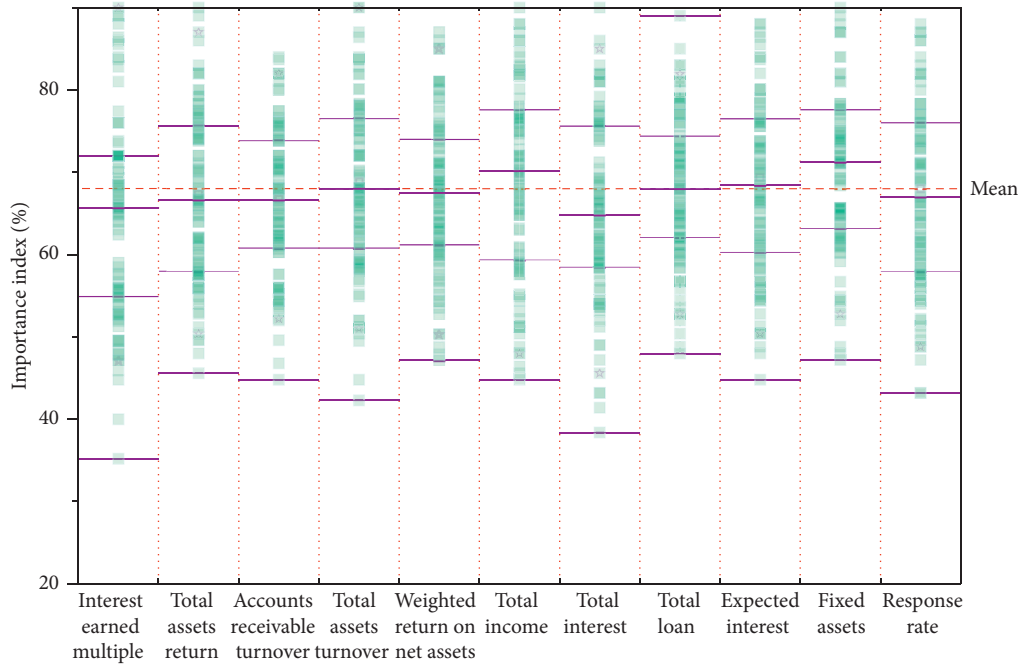
FIGURE 4: Importance of variables for assessing the overall financial risk of an enterprise.
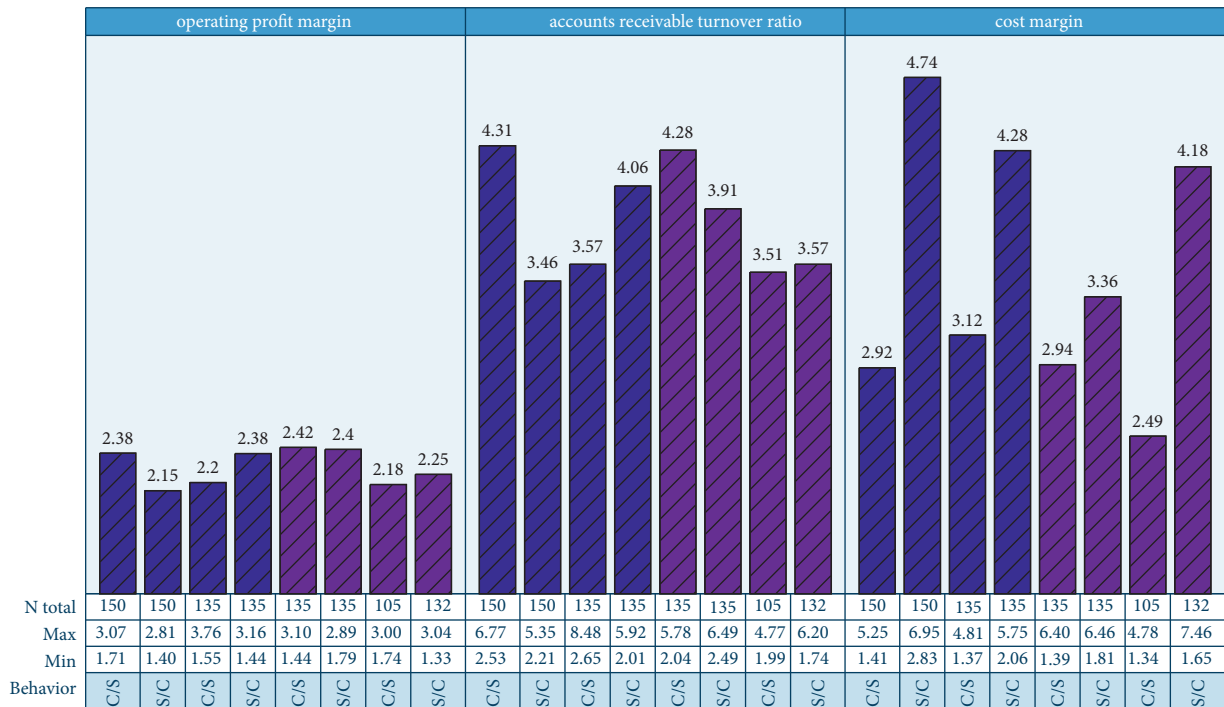
FIGURE 5: Importance of business operational risk assessment variables.

>7.552, the enterprise does not have financial risk. The interpretation of the rule set generated by the decision tree model will provide the listed company with quantitative evidence to assess the degree of the business risk of the company, which will help the listed company to measure whether it has business risk.

The importance diagram of the variables generated by the enterprise financing risk assessment model is shown in Figure 6. For the decision result of the CART decision tree model, among the three characteristics of financing risk, the importance of "interest multiples earned" becomes the most important evaluation indicator, and the importance score of

FIGURE 6: Importance of variables generated by the enterprise financing risk assessment model.

FIGURE 7: Model accuracy validation.

FIGURE 8: Loss curve of sample size set for associative memory neural network.

its indicator is close to 0.8. Therefore, the model results indicate that listed companies should focus on the following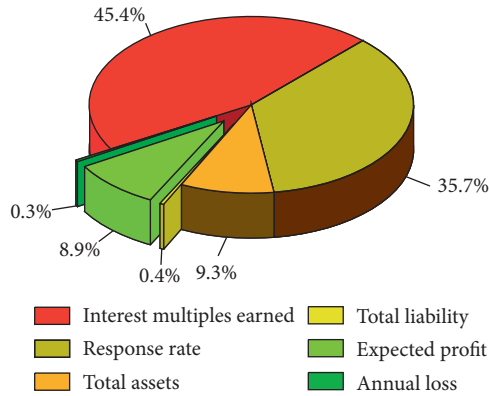 characteristics in the process of enterprise financing risk evaluation. The year-over-year revenue growth rate, net asset per share growth rate, and total asset growth rate are important indicators of a company's growth ability. The model results suggest that listed companies should focus on the indicator characteristic of "interest multiples earned" in the process of corporate financing risk assessment. The set of rules for the determination of the CART decision tree for corporate financing risk is as follows. Rule 1: when the "interest earned multiple" is ≤1.249, the enterprise has financing risk. Rule 2: when the interest earned multiple is >1.249, the firm is not at risk. The interpretation of the rule set generated by the decision tree model will provide the listed company with quantitative evidence to assess the level of funding risk of the company, which will help the listed company to measure whether it has funding risk.

From Figure 7, it can be seen that this paper's model based on associative memory neural network achieves 83% accuracy for the discriminative effect of the training set and 76% accuracy for the validation set, which is a better result. At the same time, this paper combined the GBDT model with the logistic regression model for the combined prediction, and the prediction results showed that the combined model improved the prediction accuracy of the training set to 91% and the validation set to 78%, in terms of both accuracy and stability; the combined model warning effect is more significant compared with the single model, so it proves the feasibility of the combined model. Compared with the single model logistic regression model, the ROC curve of the associative memory neural network model regression financial risk warning model is closer to the upper left axis, and its AUC value is 0.79, which is significantly higher than the logistic regression model AUC value of 0.60. The advantages of the associative memory neural network financial risk warning model can be seen directly from the ROC curve and AUC value. From the experiment, it can be concluded that GBDT for feature combination can better mine the information in the financial data of listed companies, and the logistic model has fast processing speed, which can solve the problem of slow processing speed that
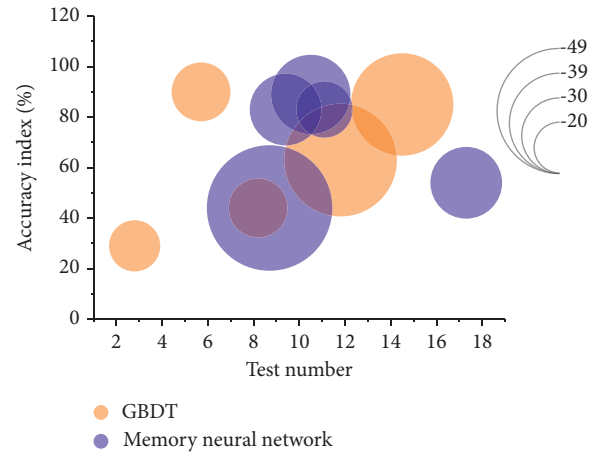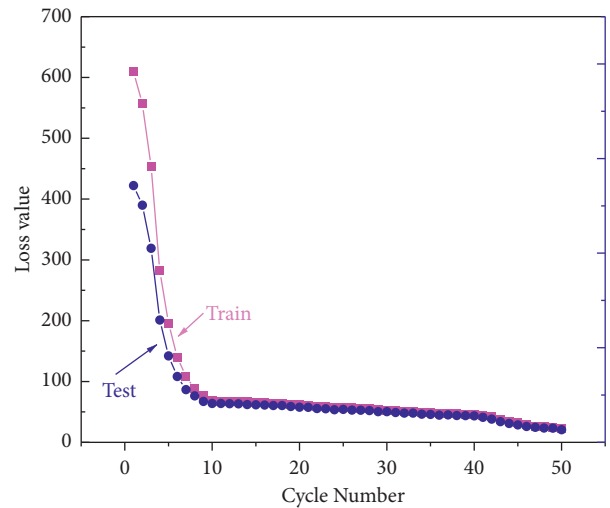
GBDT cannot be processed in parallel, and the model fused with GBDT and logistic model can be effectively used in the field of financial risk early warning of listed companies for modeling.

The learning rate is the rate at which the input variables are updated at each iteration during training and determines how far the weights have to move in the direction of the gradient in a small batch, through many iterations, eventually moving to a position that matches the training accuracy of the network. The process of learning the features of a sample dataset by an associative memory neural network is the process of constantly iterating forward. A lower learning rate makes the training process more reliable, but optimization will take longer. Higher learning rates, on the other hand, lead to nonconvergence of training and may cause very large weight changes, making the loss function very poor. Optimizing the rate of matrix multiplication and improving memory utilization can be achieved by increasing the number of batches in a certain region, which reduces the number of updates needed to complete the training of the

entire dataset and speeds up the process somewhat for data of the same capacity. Stochastic gradient descent updates only one sample information at a time, which speeds up the training, but because only one sample is used at a time, it does not represent the entire training sample, making it more difficult to converge the training results to some minimum value. For a company to achieve sustainable development, it must attach great importance to its R&D and innovation capabilities. Only through continuous innovation and continuous inventions can a company gain a competitive advantage. With the development of deep learning, it was shown that the training results could be converted to a local minimum by slowly decreasing the learning rate. After adjusting the parameters several times, the expected training results were finally obtained, and the average absolute error of the model on the samples during training finally converged. The loss curves for the sample dataset are shown in Figure 8.

The blue line in the figure represents the training set loss curve and the orange line represents the test set loss curve; the horizontal axis presents the number of iterations of the samples in the training process, and the vertical axis presents the change of the average error value during the training process. According to the parameters set in this paper, the Spyder window can display the losses of the training and test samples in each iteration and calculate the root mean square error of the model at the end of training. During the training process, the more iterations, the smaller the sampling error, with local fluctuations in a small range. At the beginning of the training, the error drops quickly, and in the training interval from 10 to 50 iterations, the error drops at a rapid rate, indicating that the model is being fine-tuned locally, and from 100 training iterations onwards, the error drops more gently, indicating that the model has converged to an optimal process. The trend of the loss curve of the training sample and the loss change of the test sample are integrated, and the error of the training sample finally converges with a better fit; the error of the test sample converges to a local minimum, and its fit is not as good as that of the training sample, but the effect of using gradient descent for optimization is obvious and will not affect the performance of the whole model.

Based on the associative memory neural network model, the early warning indicator system that can reflect the risk characteristics of listed agricultural companies is established by combining the risk characteristics of the company and the causes of its generation: it contains financial and nonfinancial indicators, covering solvency, operating capacity, development capacity, profitability, and other capacity indicators. To improve the convergence speed and stability of the model, SPSS software is used to conduct factor analysis on the data of the primary selected indicators and calculate the factor scores of the samples as the input layer data of the model. In this way, the accuracy of the financial crisis model is further determined and a part of the sample is selected to test the model. Finally, the model is applied to the actual company to identify its problems and optimize them to reduce the possibility of its financial crisis.

5. Conclusion

Associative memory neural network models are generally constructed based on data about the characteristics of the industry in question to make predictions about various aspects of demand points. Associative memory neural network model predictions are not only limited to daily life preferences, food, clothing, housing, etc. but now also almost permeate in financial industry development, financial forecasting, etc. In this paper, by analyzing the subsequent development capability in conjunction with the industry characteristics and competitive landscape of the company, it is found that the model can reflect the situation of the company well through early warning indicators and provide a reference for the future development of the company. If only the data model fitting of financial indicators is introduced, it will be too single and will not be able to judge the situation of finance comprehensively. In the study, it was found that the number of samples in the high financial risk category was too small compared to the samples in the low and medium financial risk categories, and if the sample data were not balanced, it might lead to the early warning model not learning the characteristics of the samples in the high financial risk category during training, which eventually led to a low prediction accuracy of the samples in this category. Since factors such as company composition and equity can also have an important impact on finance, this paper adds nonfinancial indicators to improve the generalizability and accuracy of the model. Nowadays, companies are in an increasingly complex market environment and social activities, and big data simulation by quantifying the relevant indicators of companies in the industry can reflect the current situation of the industry comprehensively and extensively, and the reasonable inclusion of industry-specific indicators can also improve the effectiveness of the early warning model. By combining the company indicators and the established early warning model, the principal components that can be identified by the model are used to analyze the situation in which the company risks arise while providing opinions that incorporate the details of Company B. Using this financial early warning model, one can adjust the company structure, do strategic planning and crisis prevention in advance for the changing market, organize training for financial early warning talent, strengthen the communication between financial early warning talents, and ensure a reasonable shareholding structure to avoid one share being dominant; secondly, one can improve the information construction of the financial system and establish a set of the financial system applicable to the company on this basis. At the same time, we should comply with the choice of the market, and the development of diversified markets can ensure their competition and at the same time focus on product quality.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

Acknowledgments

References

[1] M. Á. Fernández-Gámez, J. A. C. Soria, J. A. C. Santos, and D. Alaminos, "European country heterogeneity in financial distress prediction: an empirical analysis with macroeconomic and regulatory factors," *Economic Modelling*, vol. 88, pp. 398–407, 2020.

[2] M. Meenu Sreedharan, A. M. Khedr, and M. El Bannany, "A multi-layer perceptron approach to financial distress prediction with genetic algorithm," *Automatic Control and Computer Sciences*, vol. 54, no. 6, pp. 475–482, 2020.

[3] X. Yan, W. Weihan, and M. Chang, "Research on financial assets transaction prediction model based on LSTM neural network," *Neural Computing and Applications*, vol. 33, no. 1, pp. 257–270, 2021.

[4] M. Matsumaru, T. Kawanaka, S. Kaneko, and H. Katagiri, "Bankruptcy prediction for Japanese corporations using support vector machine, artificial neural network, and multivariate discriminant analysis," *International Journal of Industrial Engineering*, vol. 1, no. 1, pp. 78–96, 2019.

[5] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, "Class-imbalanced dynamic financial distress prediction based on adaboost-SVM ensemble combined with SMOTE and time weighting," *Information Fusion*, vol. 54, pp. 128–144, 2020.

[6] J. Che, S. Zhao, Y. Li, and K. Li, "Bank telemarketing forecasting model based on t-SNE-SVM," *Journal of Service Science and Management*, vol. 13, no. 3, pp. 435–448, 2020.

[7] A. Di Vaio, R. Palladino, R. Hassan, and O. Escobar, "Artificial intelligence and business models in the sustainable development goals perspective: a systematic literature review," *Journal of Business Research*, vol. 121, pp. 283–314, 2020.

[8] P. K. Mitra and C. Banga, "Predicting Indian basket crude prices through machine learning models-a comparative approach," *International Journal of Business Forecasting and Marketing Intelligence*, vol. 5, no. 3, pp. 249–266, 2019.

[9] G. Kumar, S. Jain, and U. P. Singh, "Stock market forecasting using computational intelligence: a survey," *Archives of Computational Methods in Engineering*, vol. 28, no. 3, pp. 1069–1101, 2021.

[10] K. Lu, Y. Lyu, X. Li, and Y. Zhang, "A new method for evaluating information system growth of SMEs based on improved BP neural network," *Information Systems and E-Business Management*, vol. 18, no. 4, pp. 779–792, 2020.

[11] B. S. Kumar, V. Ravi, and R. Miglani, "Predicting indian stock market using the psycho-linguistic features of financial news," *Annals of Data Science*, vol. 8, no. 3, pp. 517–558, 2021.

[12] S.-J. Lin, "Integrated artificial intelligence and visualization technique for enhanced management decision in today's turbulent business environments," *Cybernetics and Systems*, vol. 52, no. 4, pp. 274–292, 2021.

[13] S. Jones and T. Wang, "Predicting private company failure: a multi-class analysis," *Journal of International Financial Markets, Institutions and Money*, vol. 61, pp. 161–188, 2019.

[14] P. T. Thi Ngo, M. Panahi, K. Khosravi et al., "Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran," *Geoscience Frontiers*, vol. 12, no. 2, pp. 505–519, 2021.

[15] J. Liu, T. Sun, Y. Luo, S. Yang, Y. Cao, and J. Zhai, "An echo state network architecture based on quantum logic gate and its optimization," *Neurocomputing*, vol. 371, pp. 100–107, 2020.

[16] I. Bardhan, H. Chen, and E. Karahanna, "Connecting systems, data, and people: a multidisciplinary research roadmap for chronic disease management," *MIS Quarterly*, vol. 44, no. 1, pp. 185–200, 2020.

[17] M. Ding, H. Zhou, H. Xie, M. Wu, Y. Nakanishi, and R. Yokoyama, "A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting," *Neurocomputing*, vol. 365, pp. 54–61, 2019.

[18] W. R. Bittencourt and P. H. M. Albuquerque, "Evaluating company bankruptcies using causal forests," *Revista Contabilidade & Finanças*, vol. 31, no. 84, pp. 542–559, 2020.

[19] S.-J. Lin, T.-M. Chang, and M.-F. Hsu, "An emerging online business decision making architecture in a dynamic economic environment," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 2, pp. 1893–1903, 2019.

[20] K. Bhavsar, V. Shah, and S. Gopalan, "Machine learning: a software process reengineering in software development organization," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 2, pp. 4492–4500, 2020.

[21] A. N. Khan, X. Cao, and A. H. Pitafi, "Personality traits as predictor of M-payment systems: a SEM-neural networks approach," *Journal of Organizational and End User Computing*, vol. 31, no. 4, pp. 89–110, 2019.

[22] H. Naz and S. Ahuja, "Deep learning approach for diabetes prediction using PIMA Indian dataset," *Journal of Diabetes & Metabolic Disorders*, vol. 19, no. 1, pp. 391–403, 2020.

[23] W. Bao, N. Lianju, and K. Yue, "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment," *Expert Systems with Applications*, vol. 128, pp. 301–315, 2019.

Research Article

Analysis of the Evolution of Time and Space Differences and Coordinated Development Degree of Tourism Economy Based on Regional Internet of Things Technology

Su Zhang ⓘ

Tourism College of Zhejiang, Hangzhou, Zhejiang 311231, China

Correspondence should be addressed to Su Zhang; zhangsu@tourzj.edu.cn

From the time dimension, this paper analyzes the characteristics of the scale, industrial structure, employment flexibility, and comprehensive employment effects of regional tourism employment based on the three-level criteria of the scope of tourism employment based on the regional Internet of Things. From the spatial dimension, we take the regional city as the basic research unit, adopt multiple indicators, and conduct a comprehensive evaluation of the regional development of regional tourism employment through horizontal and vertical comparisons. This paper uses multiple linear regression analysis to establish the relationship between the development level of the county tourism economy and the influencing factors; in order of magnitude of influence, they are tourism resource endowment, location traffic conditions, and economic development. Using a combination of a single indicator and multiple indicators, the county tourism economy is evaluated and analyzed for differences in time and space. We select the total tourism revenue as an indicator and use methods such as range, standard deviation, coefficient of variation, and coefficient to analyze the time difference of the tourism economy in each county. We adopt the Granger causality test and other methods to analyze the factors affecting tourism employment in the area and the growth mode. Through the establishment of a structural model of the tourism employment growth dynamics system, causality test and other methods are adopted to analyze the regional tourism employment influencing factors and growth methods, and the results show that the regional tourism employment growth mode is an investment-driven tourism employment growth mode.

1. Introduction

In recent years, the development speed of the tourism industry based on the regional Internet of Things has been accelerating, and the level of development has also been increasing. Compared with other regions, especially the developed tourism regions, there is still a relatively large gap, and the overall development level is still relatively low [1]. The travel rate can reflect the number of tourists and tourism income in a region, as well as the tourism development status of a region. There are obvious temporal and spatial differences in the travel rate of cities and states in the same area [2]. In addition to the improvement of people's living standards and scientific and cultural quality, tourism has become a sunrise industry for regional economic and social development [3]. The development of the

tourism industry has greatly increased regional national income and effectively increased the gross national product and fiscal revenue; the tourism industry has a relatively large degree of relevance and a strong development driving effect, which can promote the development of related industries and the economy and promote structural optimization and adjustment [4]. The tourism industry is a labor-intensive industry that can accommodate and absorb more labor, create abundant employment space, and help solve social employment problems; the development of the tourism industry improves the quality of life of the people. Not only it can improve people's physical and mental health but also can improve people's quality and accomplishment and enrich people's spiritual world, promote the all-round development of people, and help build a harmonious society [5].

Domestic research studies on the spatial and temporal distribution characteristics of tourism flows have more research on the spatial distribution of inbound tourism flows [6]. Talari et al. [7] divided the regional inbound tourism flow into spatial divisions. For the research on the spatial pattern of tourism flow, in the horizontal pedigree, more emphasis is placed on the study of the circle structure, and in the vertical pedigree, more emphasis is placed on the single-point level study [8]. At present, the conceptual models of the spatial structure of tourism flow mainly include leisure vacation mode, travel mode, spatial layered movement mode, international tourism flow mode, urban tourism flow mode, and tourist travel spatial structure mode [9]. Piccialli and Chianese [10] conducted a research on the theory of tourism flow spatial pattern and divided the basic theory of tourism flow spatial pattern into three theories, namely, circle structure theory, core edge theory, and spatial diffusion theory. And on this basis, some problems existing in the study of tourism flow are put forward. Perles et al. [11] studied the spatial structure optimization of the regional outbound tourism flow and optimized it from the two aspects of the destination and the source of tourists by studying the characteristics of the regional outbound tourism flow and its spatial structure. Foreign scholars have mainly studied tourists' travel preferences, travel motives, travel behavior, travel demand, and the analysis and prediction of travel flows [12]. Zsarnoczky [13] analyzed five factors that directly affect the bilateral flow of tourism flows, namely, political factors, economic factors, private enterprises, intangible factors, and diplomatic and health factors. Weber et al. [14] emphasized the seasonality of tourism flows. From a macroperspective, it uses register-based data to analyze the spatial structure and analyzes the recent impact of tourist flow and migration flows and the flow of foreign property ownership. Jha et al. [15] carried out research on the comparative advantages of tourism flows. The article pointed out that tourism and international trade flows have similarities. The results are usually supported by two aspects: supply-side variables and the comparative advantages of exporting countries. In addition, western geographers are also committed to the study of the interpretation mode and spatial distribution mode of tourism flow, but they are mostly based on hypotheses, so they are of little practical significance, and they have made useful contributions to the research methods of tourism flow [16]. Some scholars analyzed five factors that directly affect the bilateral flow of tourism flows, namely, political factors, economic factors, private enterprises, intangible factors, and diplomatic and health factors [17]. Researchers emphasize the seasonality of tourism flows [18]. An article pointed out that tourism and international trade flows have similarities, and the results are usually supported by two aspects: supply-side variables and the comparative advantage of exporting countries [19]. In addition, Western geographers are also committed to the study of interpretation models and spatial distribution models of tourism flows, but most of them are based on assumptions, which are of little practical significance and have not made useful contributions to the research methods of tourism flows. The study of tourism flow is an important basis for formulating tourism development plans, determining source markets, and formulating external publicity strategies [20, 21].

At present, the research on the difference of tourism economy based on the regional Internet of Things is slightly insufficient. The existing research results are limited to large regions or intercity areas, and most of them are still qualitative analysis, lacking quantitative empirical analysis. This article uses quantitative, qualitative, and spatial analysis methods to study the spatial and temporal differences of the tourism economy based on the regional Internet of Things, analyzes the influencing factors of the differences, and then proposes countermeasures for the development of the county tourism economy on this basis. In view of the characteristics of the temporal and spatial differences in the travel rate, the reasons for these characteristics are analyzed from the aspects of politics, society, tourism resources, transportation, and short-term crises. This method of calculating how much labor input is needed to reverse the output of the tourism industry believes that the employment capacity of tourism characteristic industries mainly depends on two major factors. Based on the actual situation, suggestions are made to increase the travel rate and promote the sustainable development of the regional tourism industry: establish the concept of grand tourism and innovate the concept of tourism development, which is conducive to accelerating the system reform; strengthening the government's leadership in the tourism industry is conducive to regional tourism; focusing on the training of tourism professionals is conducive to improving the quality of tourism services.

2. Construction of a Model for the Evolution of Temporal and Spatial Differences in Tourism Economy Based on the Regional Internet of Things Technology

2.1. Regional IoT Spatial Hierarchy. In a type of real-time system such as the Internet of Things, Time Computation Tree Logic (TCTL) can represent a large number of naturally occurring normative attributes, such as security and time response, and is especially suitable for research on time reachability. Because of its powerful functions, it can capture many routine inspections and can perform verification more effectively. Since the state space of the underlying model is basically uncountable, the analysis of accessibility must use symbolic technology to describe the continuity of the clock through a finite quotient. Figure 1 shows the spatial hierarchical nesting of regional IoT.

Due to the real-time and concurrent characteristics of the Internet of Things system, based on exhaustive search for multiplying the state space and the clock band, the number of states generated tends to increase exponentially, which greatly increases time and space consumption. When the number of concurrent systems is large, searching the state space directly will exceed the memory limit of the computer, and the verification result will not be obtained.
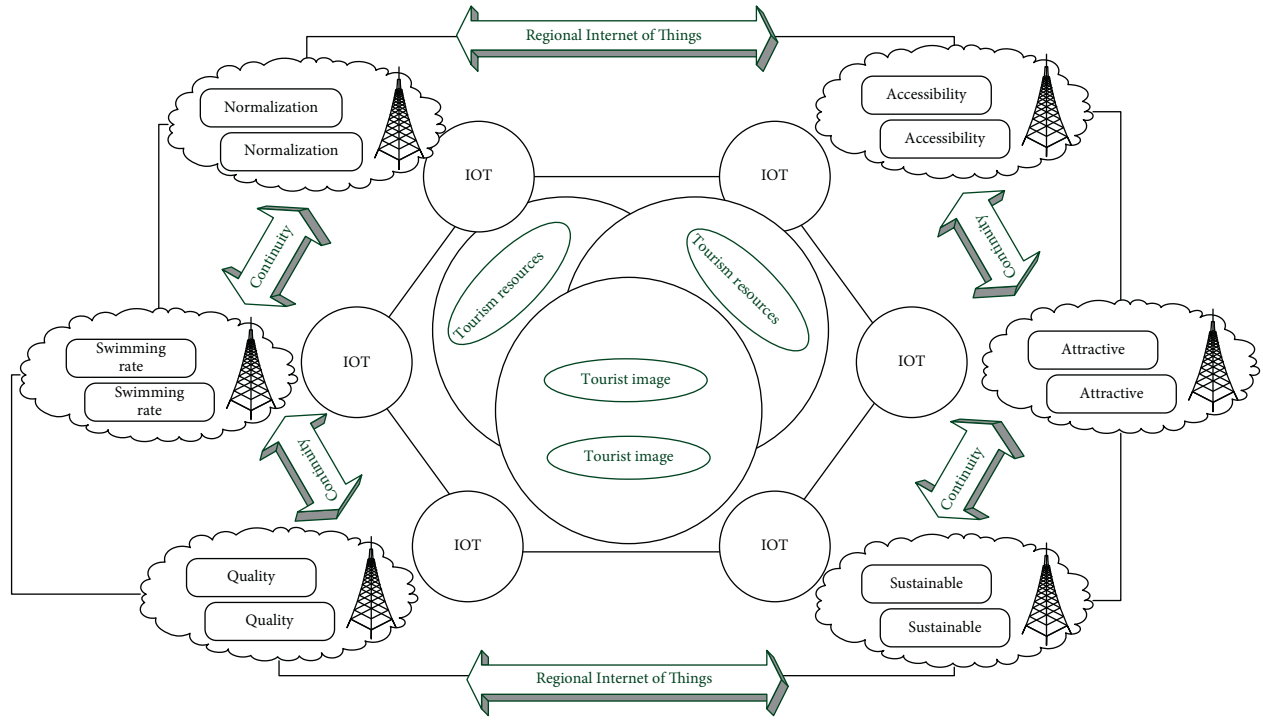
FIGURE 1: Spatial nesting of regional IoT.

$$p(x(1), x(2), \ldots, x(n)) = \prod_{i=1}^{n} p[x(n)|x(n-i+1)],$$

$$C = 2 \times \frac{(p(x) \times p(n))^{1/2}}{(p(x) + p(n))}. \tag{1}$$

The direct consumption coefficient, also called the input coefficient, is recorded as i, j, which refers to the direct consumption of the unit's total output of the j product (or industry) sector in the production and operation process. The value of the goods or services consumed in the i product sector is represented by the direct consumption coefficient of each product (or industry) sector in the form of a direct consumption coefficient table or a direct consumption coefficient matrix, usually represented by the letter T.

$$f(x) = x(i) + \frac{x(i) \times x(j)}{x(i) + x(j)},$$

$$T^2 - \frac{1}{n} \times \sum_{i=1}^{n} (p(i) - p(x))^2 = 0. \tag{2}$$

The direct consumption coefficient reflects the basic characteristics of the production structure in the Leontief model and is the basis for calculating the complete consumption coefficient. It fully reveals the technical and economic connections between the various sectors of the national economy, that is, the strength of the interdependence and mutual restraint between the sectors, and provides important economic parameters for constructing input-output models.

$$\begin{cases} x(i,0) = a - x(j,0), \\[2mm] x(i,1) = b - x(j,1), \\[2mm] x(i,j) = \sqrt{\dfrac{x(j,i)}{a+b}}. \end{cases}$$

$$U(x) = (I - C)^{-1} \times C^{T} - \sum_{i=1}^{n} p(x) \times C(i) \sum_{i=1}^{n} p(x) \times Z(i). \tag{3}$$

Range refers to the difference between the maximum value and the minimum value of a certain index. It reflects the largest absolute difference in the change of a certain indicator among various regions. The range index g can indicate the gap between the county with the most developed tourism economy and the county with the least developed tourism economy. The standard deviation is the square root of the square root of the square of the deviation of the standard value of each unit of the population from its mean.

$$g(w(i)|w(i-n+1), \ldots, w(i-1)) = \frac{P(w(i-n+1), \ldots, w(i-1), w(i))}{P(w(i-n+1), \ldots, w(i-1))}. \tag{4}$$

The complete consumption coefficient refers to the sum of the direct and indirect consumption of goods or services in the i product sector for each unit provided by the j product sector for final use. The complete consumption coefficient of each product department is expressed in the form of a table, which is a complete consumption coefficient table or a complete consumption coefficient matrix, usually represented by the letter V.

$$V(x) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w(i,j) * (x(i) - x) \times (x(j) - x)}{\sum_{i=1}^{n} \sum_{j=1}^{n} w(i,j) \times x(i,j)}. \quad (5)$$

The coefficient of variation is the ratio of the standard deviation to the corresponding average. It can exclude the influence of the difference in units or averages on the comparison of the degree of variation of two or more values. The coefficient of variation $f(x)$ can explain the relative differences in tourism economy among different research units.

$$\exp \sum_{i=1}^{n} x(i) f(s,t) = \exp x(1) f(x) + \cdots + \exp x(n) f(x). \quad (6)$$

2.2. Spatiotemporal Difference Evolution Model.

In this article, the difference in the tourism economy based on the regional Internet of Things means that, under the influence of different factors, there are differences in economic levels between different regions in the development of the tourism economy and the unevenness in the development of tourism industry in different regions within a certain period of time. Figure 2 shows a schematic diagram of the evolution model of temporal and spatial differences. The growth pole refers to a collection of economic units with agglomeration characteristics in space. Economic development will first explode at a certain point in a region and then gradually spread outward from this point, thereby driving the economic development of the entire region. It has a strong economy of scale and it drives the economic development of the entire region through multiplier effects, dominance effects, polarization, and diffusion effects. Some areas developed first to become regional "centers" due to various favorable factors, while other areas became "peripheries" due to lagging development. The center is in a dominant position, while the periphery depends on the center for development, and peripheral funds, materials, and labor will be continuously transferred to the center.

Generally speaking, the lower the level of tourism development is, the lower the consumption level of tourists is, and the higher the proportion of basic tourism consumption expenditure in total consumption expenditure is. At present, as the regional tourism industry is generally in the early stage of development, tourism supporting facilities and services are not yet complete, and the products are single, tourism consumption is mainly used for transportation, accommodation, catering, sightseeing, etc., and the proportion of nonbasic tourism consumption is very small. The change curve of employment elasticity is basically similar to that of the tertiary industry, and the fluctuation range shows a trend of gradual convergence. The employment elasticity of tourism industries showed almost the same changes as the employment elasticity. The purpose of statistical testing is to test the reliability and accuracy of estimated values. Table 1 shows the statistical test results of tourism characteristic industries. Goodness of fit refers to the degree of fit of the regression line to the observed value. The regression parameter estimation shows that the model fits the sample well.

The tourism characteristic industry belongs to the tertiary industry. There are two main methods for measuring the difference of the tourism economy: relative difference and absolute difference. It shows that the regression equation is significant; that is, the combination of variables such as tourism resource endowment, location and transportation conditions, economic development level, and industrial structure does have a significant impact on the tourism economy. When other variables remain unchanged, the explanatory variables such as the endowment of tourism resources, location and transportation conditions, and the level of economic development have a significant impact on the tourism economy, while the industrial structure has no significant impact on the tourism economy.

2.3. The Distribution of Tourism Economic Weighting Factors.

The establishment of an indicator system of factors influencing the tourism economy based on the regional Internet of Things requires that the indicators have scientific and clear characteristics and can accurately reflect the main aspects and essential characteristics of factors affecting tourism economy in the county. Based on the correct understanding of the county tourism economy, the influencing factors are classified, and each category corresponds to specific indicators so as to construct the county tourism economy influencing factor index system. Therefore, it is of great significance to develop the tourism industry, to further utilize the potential of tourism characteristic industries to absorb labor, to adjust the regional employment structure, to optimize and upgrade the employment structure, and to achieve the coordinated development of the employment structure and industrial upgrading.

Based on the proximity of counties (cities and districts), the binary adjacency weight matrix is used, and Geoda software is used to calculate the global spatial autocorrelation coefficient of the comprehensive development level of the regional tourism economy. The adjusted goodness of fit was 0.918, indicating that there is a significant linear relationship between the dependent variable and the independent variable. The statistical t values are all positive, and the significance level is less than 0.05, indicating that the standard regression coefficient is significant. Figure 3 shows the two-dimensional scatter point distribution of the goodness of fit of the temporal and spatial differences. Therefore, the equation has a good fitting effect on the whole, indicating that the level of regional economic development can reflect the development of the regional tourism economy to a certain extent. Therefore, when constructing indicators for the comprehensive development level of the tourism economy, this article constructs it from two aspects: tourism
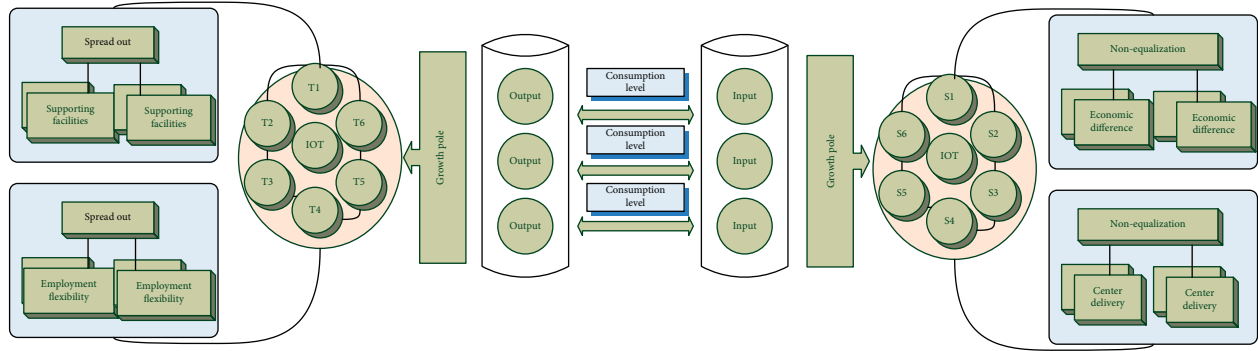
FIGURE 2: Schematic diagram of the evolution model of temporal and spatial differences.

TABLE 1: Statistical test results of tourism characteristic industries.

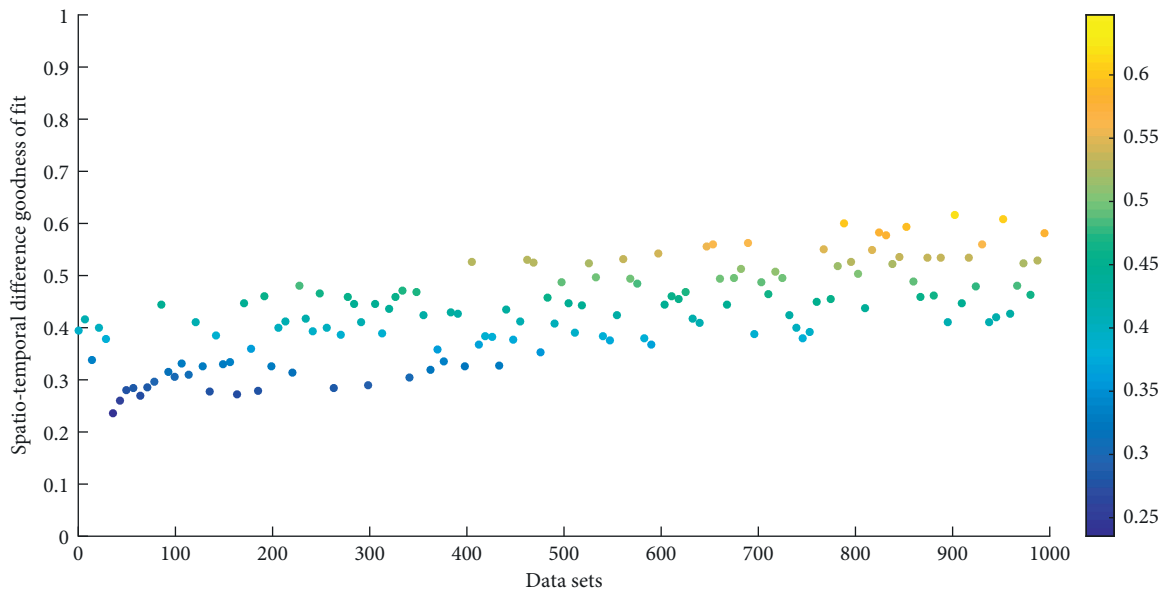| Unit index	Coefficient of elasticity	Goodness of fit	Return parameters
1	0.03	0.97	1.44
2	0.07	0.96	3.71
3	0.05	0.89	2.36
4	0.08	0.87	2.53

FIGURE 3: Two-dimensional scatter point distribution of goodness of fit of temporal and spatial differences.

industry development level and regional economic development level.

The selected index system needs to be able to scientifically reflect the actual situation of the county tourism economy and faithfully reflect the connotation of the corresponding influencing factors by selecting indicators with strong relevance and high representativeness. The establishment of the index system is ultimately to implement each specific data, so the availability and calculation of the selected index data are required to be strong. Because it is relatively difficult to collect county data, it is necessary to ensure the availability of indicators. The principle of operability is one of the indispensable conditions for establishing a complete index system. The full consumption coefficient refers to the sum of the direct and indirect consumption of goods or services in the i-th product sector for each unit provided by the j-th product sector for final use. Based on the characteristics of county tourism economy and the principles of influencing factor index system construction, and drawing on some of the latest results of current regional tourism economic research, this article focuses on the four aspects of tourism resource endowment, location and transportation conditions, economic development level, and industrial structure. They are tourism resource endowment, location traffic conditions, economic development level, and industrial structure. Therefore, we can

conclude that the endowment of tourism resources affects the tourism economy of the county to a greater extent. The location and traffic conditions are greater than the level of economic development and the industrial structure. This is basically consistent with empirical judgment.

3. Results and Analysis

3.1. Extraction of Tourism Economic Data. We use SPSS19.0 software to analyze the selected indicators using the orthogonal rotation method with maximum variance. The KMO test result was 0.819, indicating that the principal component analysis method is suitable. Mathematical statistics cannot explore the evolution of economic differences from a spatial perspective. Therefore, this article uses the natural fracture method in ArcGIS software to classify counties and visualize them to analyze the development of the county tourism economy in terms of time. The characteristics of spatial evolution are to make up for the limitations of mathematical statistics in spatial analysis and also reflect the characteristics of geographic research—analyzing the differences in the distribution of things from a spatial perspective. The research needs to use the relevant data of county-level administrative districts. Table 2 shows the statistics of tourism economic data. This article is to obtain the relevant data in a complete and comprehensive way. By visiting statistical websites of various counties, relevant effective data are collected from statistical yearbooks, statistical reports, or annual national economic and social development reports to reflect the county's tourism economy as comprehensively and accurately as possible.

Through calculations, in the past ten years, the employment scale of the regional tourism industry has shown an overall growth trend, but the volatility is relatively large. The number of employees in the regional tourism industry has reached 961,300, an increase of 24.31% compared to 773,300, and an average annual growth of 2.03%. The proportion of employees in tourism characteristic industries in the total number of employees in the society has increased, from 3.20% to 3.43%, but the proportion of employees in the tertiary industry has declined, from 14.45% to 11.35%. The growth rate of employment is faster than the growth rate of the number of employees in the whole society but obviously lags behind the growth of the number of employees in the tertiary industry. Traditional labor-intensive industries such as the postal industry have become the main industries for absorbing labor in the tertiary industry in the region, with an employment ratio of more than 10%. It can be seen that the selected index data reflecting the level of regional tourism economic development have been subjected to principal component analysis and dimensionality reduction, and four principal components have been extracted. The characteristic value of principal component one is 5.665, and its principal component contribution rate is 47.206%. The characteristic value of principal component two is 1.737, and its principal component contribution rate is 14.472%. The characteristic value of principal component three is 1.319, and its principal component contribution rate is 10.992%. The characteristic value of principal component

four is 1.178, and its principal component contribution rate is 9.819%. Its cumulative variance contribution rate has reached 8.489%. Figure 4 shows the statistical distribution of the contribution rate of tourism economic variance. The eigenvalues of the four principal components are all greater than 1, indicating that the extracted four principal components can well describe the comprehensive difference of tourism economic differences.

We use multiple linear regression to analyze the correlation between tourism economic development level and tourism resource endowment, location traffic conditions, economic development level, industrial structure, and other factors and judge the degree of influence of different factors on dependent variables. From the analysis of the causality test results of the two variable combinations composed of each influencing variable and the employment elasticity of tourism characteristic industries, there is a certain one-way relationship between labor factors and the employment elasticity of tourism characteristic industries. Both the differences between prefectures and cities and the differences within prefectures and cities show a downward trend in fluctuations. The decline from 0.934 to 0.469 shows that the regional differences in total tourism revenues within cities are gradually decreasing.

4. Simulation of Temporal and Spatial Difference Evolution Model

In order to further comprehensively study the regional differences in regional tourism, this article extracts $X1$ domestic tourism income, $X2$ tourism foreign exchange income, $X3$ domestic tourists, and $X4$ international tourists. As an explanatory variable of the effect of regional tourism employment, it is not only necessary to analyze it from the scale of employment but also to grasp it from the quality of employment. Therefore, the absolute amount of tourism employment and the elasticity of tourism employment are selected for measurement. Here, GH, TH, GT, and TT are used to denote the employment scale of tourism core industry, the employment elasticity of tourism core industry, the employment scale of tourism characteristic industries, and the employment elasticity of tourism characteristic industries. Such a statistical caliber is obviously not enough to reflect the actual situation and overall picture of the tourism industry, and it cannot reflect the corresponding status and role of the tourism industry in the national economy. Figure 5 shows a line chart of the proportion of the tourism economy in the total value. This article uses international and domestic tourist visits (TP) and international and domestic tourism revenue (TC) to characterize tourism demand factors, selects fixed asset investment (FAI) and tourism institutions (TS) in industries that are highly related to tourism economic growth to characterize tourism supply factors, and uses gross domestic product (GDP), the proportion of tertiary industry gross product (TIG), and per capita income (PCI) to characterize economic environmental factors.

It is difficult to select appropriate variables for the influencing factors of county tourism spatial differences. In

TABLE 2: Tourism economic data statistics.

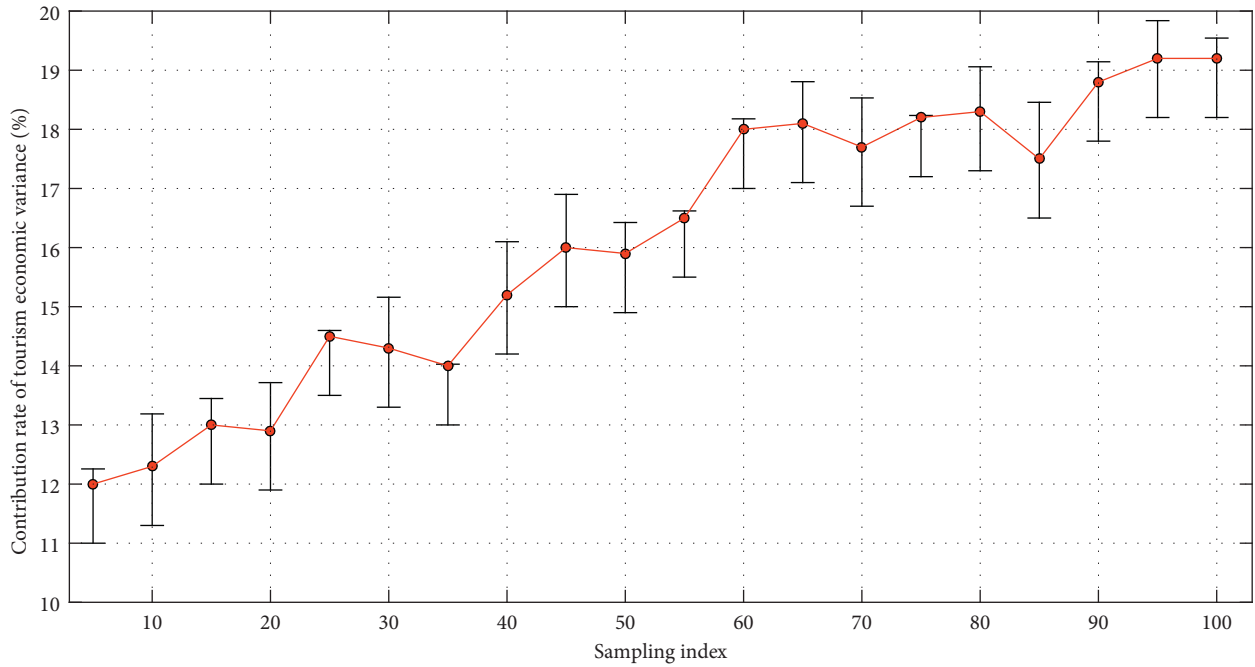| Area code	Employed population ($\times 10^4$)	Growth rate (%)	Variance contribution rate (%)
1	91.13	22.13	37.56
2	64.71	18.52	39.07
3	82.02	19.17	18.84
4	77.52	26.03	24.57
5	79.72	9.42	47.06

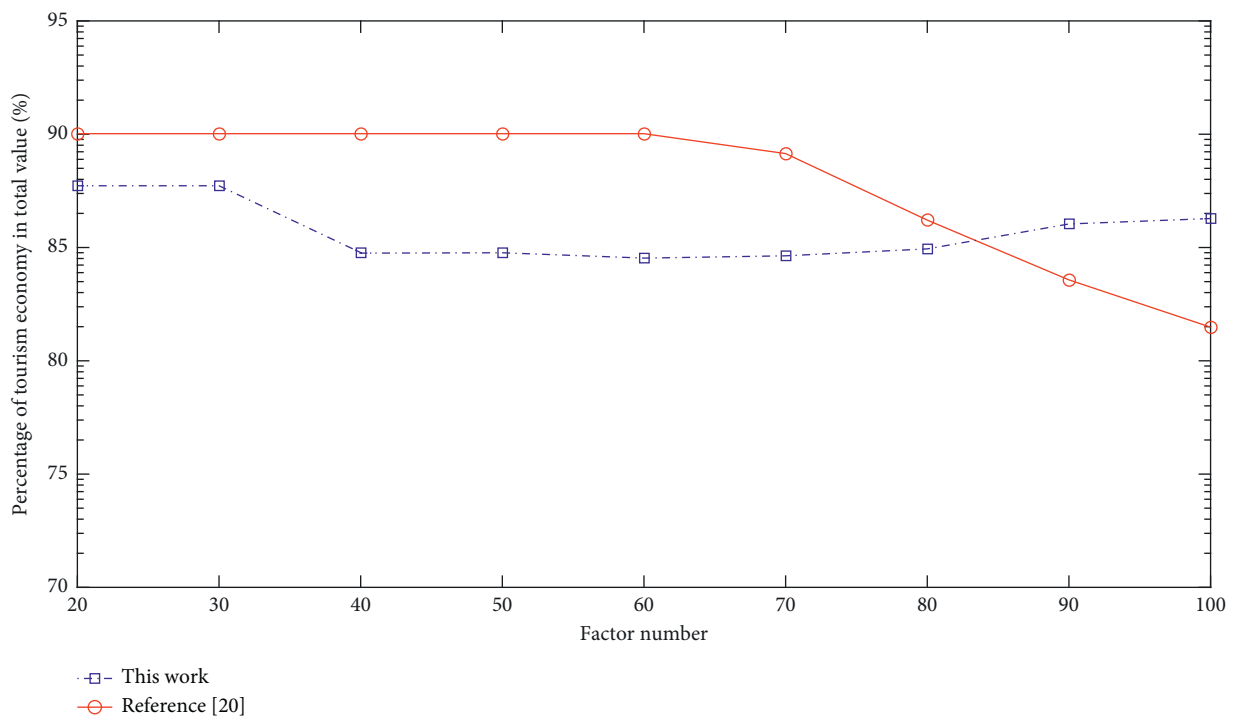FIGURE 4: The statistical distribution of the contribution rate of tourism economy variance.

FIGURE 5: Line chart of the proportion of tourism economy in the total value.

order to better measure each influencing factor, it is necessary to rationally sort and process the original data. This article has made many adjustments to the selection of data through multiple calculations of SPSS. The contribution index of the county tourism economy to the national economy is the proportion of the county's total tourism revenue to the regional GDP. According to the calculation results, the county tourism industry with a contribution index between 0 and 0.10 is named a general industry, the county tourism industry with a contribution index between 0.10 and 0.20 (including 0.10) is named as a pillar industry, and the county-level tourism industry with a contribution index between 0.20 and 0.30 (including 0.20) is named an important pillar industry, and the contribution index is at 0.20. County tourism industries above 30 (including 0.30) are named as strategic pillar industries. Figure 6 shows a fan chart based on the contribution index of the regional Internet of Things. And through ArcGIS software, these four levels of counties are presented in different colors to analyze the spatiotemporal evolution analysis of the difference in the contribution of the county tourism economy to the national economy. The total tourism income of the counties (cities and districts) in the district is extremely poor, and the standard deviation is increasing. The coefficient of variation and Theil coefficient show a downward trend in fluctuations, indicating the absolute value of the total tourism income of the counties (cities and districts) in the district. The difference shows an accelerated growth trend, and the relative difference in total tourism revenue is gradually shrinking.

We divide the total swimming rate into three levels: in the first level, the total swimming rate is extremely low, which is less than 100%; in the second level, the total swimming rate is low, which is 100% to 200%; and in the third level, the general swimming rate is high, which is between 200% and 300%. We can see that the travel rate anomalies of the cities and states in the region are obvious differences in the value. In some areas, the anomaly values are positive and some are negative. According to the data, we divide the total travel rate anomalies of each city and state into three categories: in the first category, the total travel rate has an anomaly value greater than 50%; the second category refers to areas where the total travel rate is close to the average level, and the total travel rate anomaly is lower than 50%; and the third category refers to areas with a large proportion of total employment in tourism, tourism is also the main area for absorbing labor in the region, accounting for a large proportion of total employment in the region, and the proportions of the cities and districts are basically the same. The inbound travel rate anomaly is divided into two categories: the first category is the inbound travel rate level and the second category is areas where the inbound travel rate is lower than the average level, and the anomaly value is negative. The global index of the total tourism revenue of each county (city and district) in the region is positive and shows a trend of increasing in fluctuation, indicating that there is a significant spatial correlation in total tourism revenue.
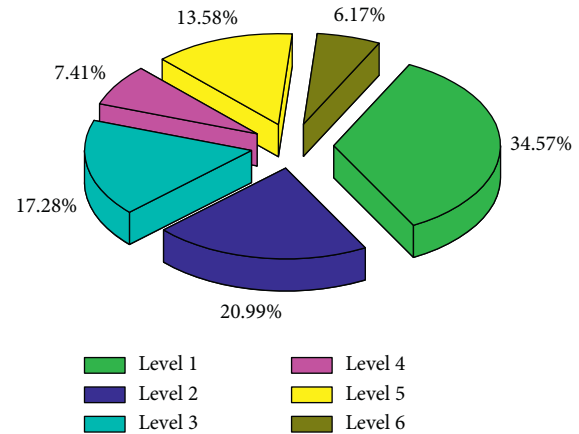

FIGURE 6: Fan chart of contribution index based on regional IoT.

4.1. Example Application and Analysis. We divide the inbound travel rate into four levels. The first level refers to areas with an inbound travel rate of less than 0.1%, the second level refers to areas with an inbound travel rate of 0.1% to 1%, the third level refers to areas where the inbound travel rate is between 1% and 2%, and the fourth level refers to areas where the inbound travel rate is greater than 2%. By comparing the number of tourism employment driven by unit income, it reflects the contribution of tourism to the employment of various regions and cities in the region and at the same time measures the imbalance in the development of this indicator between regions. Its size depends on the growth rate of the tourism economy and labor productivity. From a spatial comparison, the better the economic foundation is and the stronger the comprehensive strength of tourism economy development is, the smaller the number of jobs driven by tourism income is, basically between 0.10 and 0.15. Both the extreme value ratio and the coefficient of variation show an obvious growth trend, which can be seen intuitively. Figure 7 shows the histogram of the average annual growth rate of the tourism economy. From the comparison of the number of tourism employment driven by tourism income, the tertiary industry GDP, and the unit income of the regional GDP, the number of employment driven by tourism income per labor is the smallest. These reflect that, from the perspective of industry comparison, the output level of unit tourism labor is relatively high; from the perspective of prefecture and city comparison, the tourism labor productivity of economically developed regions is relatively high. In other words, the tourism economy of underdeveloped regions has a relatively high strong labor-absorptive capacity.

Counties with a ranking improvement of more than 10 places are named rapid development type, and counties with a ranking improvement between 0 and 10 places (including 0 and 10) are named development type, and counties with ranking regressed by 1 to 10 places (including 10) are named lagging type, and counties with ranking regressed by more than 10 places are named significantly lagging type. And we use ArcGIS software to make a spatial visualization map to explore the law of spatial change. The scale of fixed asset
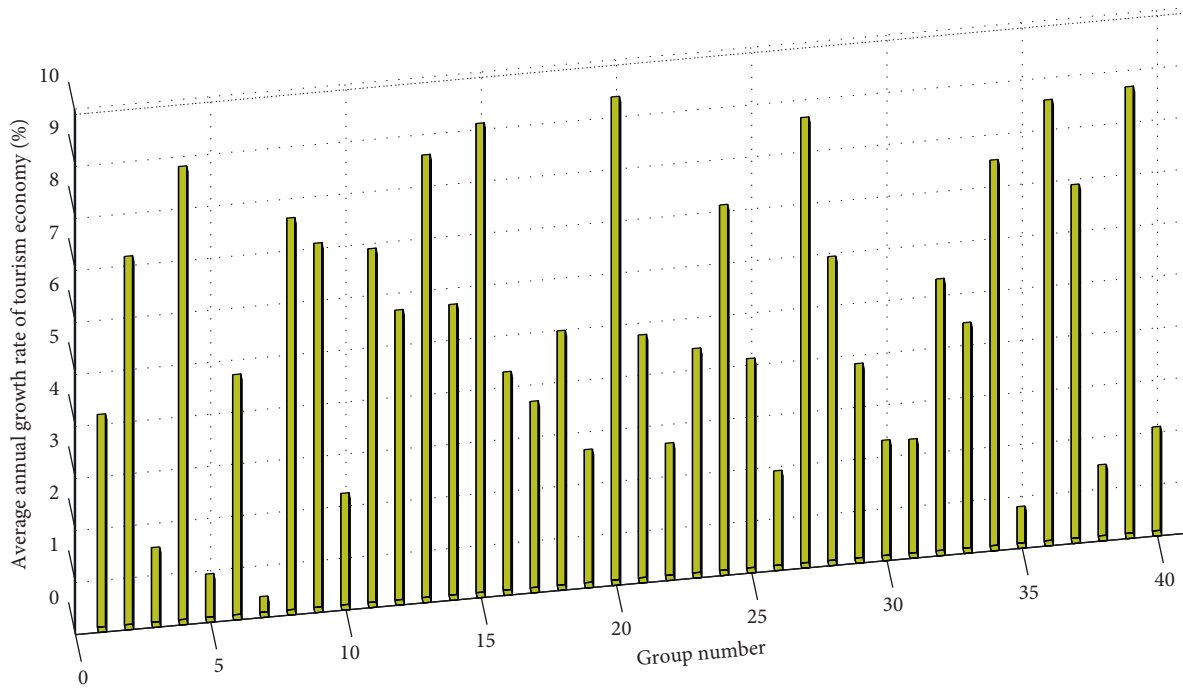
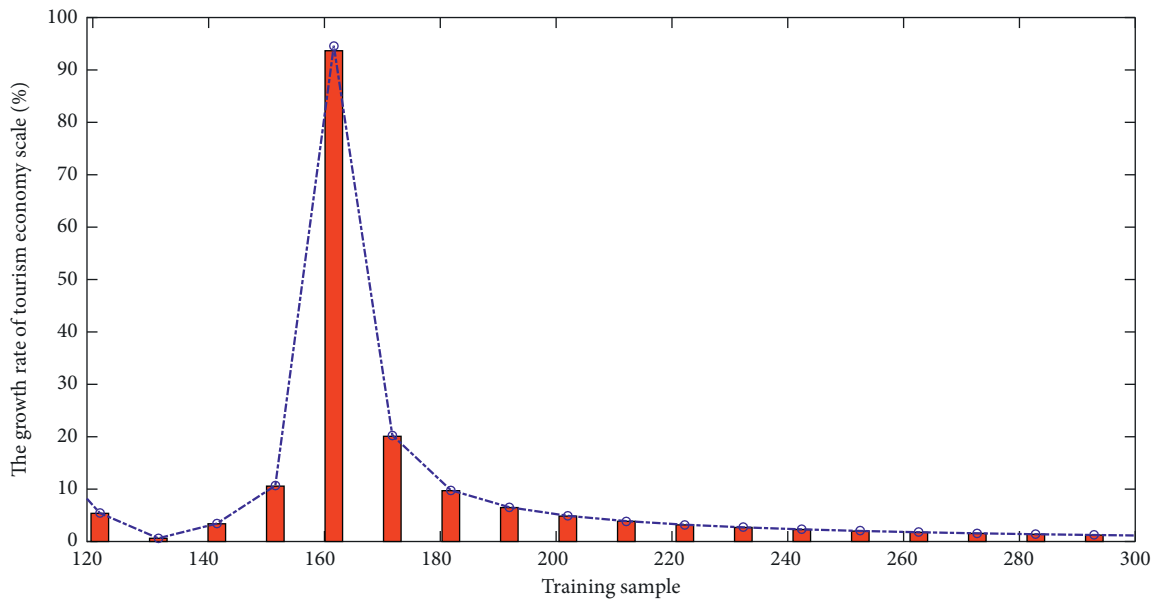FIGURE 7: Histogram of average annual growth rate of tourism economy.

FIGURE 8: Line chart of growth rate of tourism economy scale.

investment in the region has shown the characteristics of continuous increase, with an average annual growth rate of 18.25%, which is nearly 5 percentage points higher than GDP. The contribution rate of fixed asset investment to economic growth has reached 58.40%, which has risen to around 70% in recent years. Fixed asset investment has the dual function of regulating supply and demand. The long-term growth of accumulation and investment is the most basic condition for a large amount of surplus labor to be absorbed by the economy. On the one hand, continuous

large-scale investment balances the original labor and capital. The relationship is broken, which promotes employment growth; on the other hand, it indirectly drives the expansion of investment and employment in other related sectors through the multiplier effect. The range increased from 149.54 to 718.83, and the average annual growth rate reached 20.4%. And in recent years, the growth has been particularly obvious, and the range has grown rapidly from 483.87 to 718.83, a growth rate of 48.56%. The growth of the standard deviation is very similar to that of the extreme

deviation, but the growth rate is faster. Figure 8 shows a line chart of the growth rate of the tourism economy.

Through the analysis of the correlation between regional tourism employment effect variables and various influencing factors, it can basically be judged that the employment scale of core tourism industries and characteristic industries has a strong same-directional relationship with various influencing factors, and the significance level is obvious. The employment elasticity of characteristic industries has a weak and inverse relationship with various influencing factors, and its significance level is not obvious. It can be found that regional tourism can have an impact on almost all industrial sectors in terms of employment, but the employment effects of different industrial sectors are affected by tourism. The focus of this method is to explore the distribution of spatial data, so it is necessary to measure the correlation coefficient between the data. According to a comprehensive analysis, the accommodation and catering industries are most affected. The comprehensive employment coefficient of each tourism characteristic industry is 2.0015, that is, every characteristic industry gains 10,000 yuan of added value, which can lead to 20,015 employment opportunities in the accommodation and catering industry. This is followed by the manufacturing industry at 1.9477 and the tourism core industry at 1.4686.

5. Conclusion

This paper analyzes the statistical data of tourism economy based on the regional Internet of Things and concludes that the absolute difference in the unbalanced development of the county tourism economy is expanding, the relative difference is constantly shrinking, and the degree of concentration of the county tourism economy is decreasing, and it is developing in a balanced direction. This article analyzes the temporal and spatial evolution characteristics of county tourism economic differences from three aspects: the difference in the total county tourism income, the difference in the contribution of the county tourism economy to the national economy, and the difference in the ranking changes of the county tourism economy in the province. This paper follows the structural model of the tourism development dynamic system, combines the theoretical ideas of employment growth, considers the unique attributes of the tourism industry, and establishes an interactive type composed of tourism demand and tourism element supply and connected by the economic environment and labor conditions. The focus of this method is to explore the distribution law of spatial data, so it is necessary to measure the correlation coefficient between the data to determine whether the data has a spatial relationship and the strength of the relationship. From the spatial dimension, we take the regional city as the basic research unit, take the number of employment in tourism characteristic industries as the research caliber, adopt multiple indicators, and conduct a comprehensive evaluation of the regional development of regional tourism employment through horizontal and vertical comparisons. We use SPSS19.0 to establish a multiple linear regression model of important factors affecting regional tourism economy to analyze the factors affecting regional tourism economic differences. Finally, the countermeasures and suggestions for promoting the sustainable development of the regional tourism economy are put forward, with a view to narrowing the difference of the regional tourism economy and promoting the coordinated development of the tourism economy.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] Y. Lu, S. Papagiannidis, and E. Alamanos, "Internet of Things: a systematic review of the business literature from the user and organisational perspectives," *Technological Forecasting and Social Change*, vol. 136, pp. 285–297, 2018.

[2] W. Zhou, "Economic marketing of marine leisure tourism based on internet of things," *Journal of Coastal Research*, vol. 112, pp. 87–89, 2020.

[3] X. Li, "Research on tourism industrial cluster and information platform based on Internet of things technology," *International Journal of Distributed Sensor Networks*, vol. 15, no. 7, p. 155, 2019.

[4] W. Wang, N. Kumar, J. Chen et al., "Realizing the potential of the internet of things for smart tourism with 5G and AI," *IEEE Network*, vol. 34, no. 6, pp. 295–301, 2020.

[5] C. Zhang, "Design and application of fog computing and Internet of Things service platform for smart city," *Future Generation Computer Systems*, vol. 112, pp. 630–640, 2020.

[6] J. M. Tien, "Internet of things, real-time decision making, and artificial intelligence," *Annals of Data Science*, vol. 4, no. 2, pp. 149–178, 2017.

[7] S. Talari, M. Shafie-Khah, P. Siano, V. Loia, A. Tommasetti, and J. Catalão, "A review of smart cities based on the internet of things concept," *Energies*, vol. 10, no. 4, p. 421, 2017.

[8] B. P. Gautam, H. Asami, and A. Batajoo, "Regional revival through iot enabled smart tourism process framework (STPF): a proposal," *Advanced Intelligent Systems*, vol. 9, pp. 743–748, 2016.

[9] D. Shin, "A living lab as socio-technical ecosystem: Evaluating the Korean living lab of internet of things," *Government Information Quarterly*, vol. 36, no. 2, pp. 264–275, 2019.

[10] F. Piccialli and A. Chianese, "The internet of things supporting context-aware computing: a cultural heritage case study," *Mobile Networks and Applications*, vol. 22, no. 2, pp. 332–343, 2017.

[11] A. Perles, E. Pérez-Marín, R. Mercado et al., "An energy-efficient internet of things (IoT) architecture for preventive conservation of cultural heritage," *Future Generation Computer Systems*, vol. 81, pp. 566–581, 2018.

[12] Y. Ma, Y. Wang, and J. Yang, "Big health application system based on health internet of things and big data," *IEEE Access*, vol. 5, pp. 7885–7897, 2016.

[13] M. Zsarnoczky, "How does artificial intelligence affect the tourism industry?" *VADYBA*, vol. 31, no. 2, pp. 85–90, 2017.

[14] M. Weber, D. Lučić, and I. Lovrek, "Internet of Things context of the smart city," *Smart Systems and Technologies*, vol. 4, pp. 187–193, 2017.

[15] S. Jha, R. Kumar, and J. M. Chatterjee, "Collaborative handshaking approaches between internet of computing and internet of things towards a smart world: a review from 2009–2017," *Telecommunication Systems*, vol. 70, no. 4, pp. 617–634, 2019.

[16] B. B. Gupta and M. Quamara, "An overview of Internet of Things (IoT): Architectural aspects, challenges, and protocols," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 21, p. 4946, 2020.

[17] Y. Liu, M. Peng, G. Shou, Y. Chen, and S. Chen, "Toward edge intelligence: Multiaccess edge computing for 5G and Internet of Things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6722–6747, 2020.

[18] M. Zarei, A. Mohammadian, and R. Ghasemi, "Internet of things in industries: a survey for sustainable development," *International Journal of Innovation and Sustainable Development*, vol. 10, no. 4, pp. 419–442, 2016.

[19] A. Sinha, P. Kumar, and N. P. Rana, "Impact of internet of things (IoT) in disaster management: a task-technology fit perspective," *Annals of Operations Research*, vol. 283, no. 1, pp. 759–794, 2019.

[20] R. Girau, M. Anedda, and M. Fadda, "Coastal monitoring system based on social Internet of Things platform," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1260–1272, 2019.

[21] D. Li, L. Deng, and Z. Cai, "Statistical analysis of tourist flow in tourist spots based on big data platform and DA-HKRVM algorithms," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 87–101, 2020.

Research Article

A Study of Piano Timbre Teaching in the Context of Artificial Intelligence Interaction

Cui Wei (ID)

Academy of Music, Hebei Normal University, Shijiazhuang, Hebei 050000, China

Correspondence should be addressed to Cui Wei; cuiwei@hebtu.edu.cn

This paper provides an in-depth analysis and research on piano timbre teaching in the context of artificial intelligence interaction, a bold vision of piano teaching, proposes a feasible solution in terms of teaching modules in intelligent piano teaching for senior teachers, and proposes an implementation path for the integration of intelligent piano and piano teaching from the four main blocks of piano teaching. Based on the multiplicative harmonic model of monophonic signal, combined with the variability of timbre characteristics, an audio synthesis model with editable timbre is proposed, and the experimental results show that editing the timbre parameters in the model can realize timbre modification, and the synthesized timbre conforms to the piano timbre characteristics. Based on the timbre analysis and the timbre synthesis model, a piano timbre library generation system is designed. The detailed design of the software modules such as audio file reading and writing, audio information analysis, timbre parameter acquisition, timbre synthesis, and simulated performance is given. The system can generate piano timbre libraries of different qualities flexibly and meet the requirements of timbre realism. The teaching experiment designed for teaching practice from solo teaching, and the practice target is first-year undergraduate students in the university, and the practice period is six weeks, and finally, the feasibility of intelligent piano teaching application is analysed by combining the experimental results. Through the teaching objectives, teaching content, and teaching methods, teaching environment reflects intelligent piano teaching to make up for the limitations of traditional piano teaching. Analyse the development trend of intelligent piano teaching in the context of artificial intelligence interaction, and explore the value of intelligent piano teaching.

1. Introduction

With the continuous development of contemporary art, the combination of musical art expression and modern computer technology is becoming increasingly widely used. Analysing the situation of major international musical instrument exhibitions in recent years, the musical instrument manufacturing industry in developed countries has started the research and development of intelligent pianos and has made great progress. Some internationally renowned piano brands have started to display sample pianos equipped with modern technological functions, which gradually attract the attention of the international music and musical instrument industry [1]. The imported audio file needs to be parsed to obtain audio data information. Then, the audio data is frequency spectrum transformed, and the timbre characteristic information is analysed; the timbre parameters are obtained through the human-computer interaction interface, and the piano timbre is generated. Play the generated timbre in simulation, if the timbre effect is satisfactory. Some experts predict that intelligent pianos will be the direction of development in the piano industry. For digital pianos and smart pianos, the sound source is an important factor that affects their quality and grade. An intelligent piano with high-quality sound sources and characteristics can form good and continuous competitiveness in the market [2]. The intelligent digital piano series launched by Helen Piano has very high requirements for the sound source, taking the Helen Intelligent Piano Angel Series DUAII model as an example, it adopts a dual sound source system, a high-quality digital electroacoustic sound source, and a high-quality soundboard physical sound source, which is the same as the traditional piano sound, pure sound collection of 9 kinds of famous European piano quality tones, up to 7 layers

of tone performance, 128 realistic tones [3]. With the tone synthesis system to simulate the real tone of the instrument and the flexibility to modify and create new tones by editing the tone, it can enable smart and digital instruments to obtain a richer expression of the instrument. In the field of art, visual art and auditory art are the most common forms of art, each of which can bring both physical and psychological aesthetic experiences to people [4]. Traditionally, auditory art can only be experienced through the auditory senses, and the form of experience is relatively single.

With the development of multimedia technology, people put forward higher requirements for the aesthetic experience of auditory art and seek more diverse forms of artistic expression. Taking the theory of wisdom education as a guide and exploring improvement countermeasures based on reflecting on the problems in today's intelligent piano-assisted basic piano teaching in colleges and universities are both a deepening of the theory of basic piano teaching and an application and deepening of the theory of wisdom education in music education [5]. This study helps to clarify the status of intelligent piano-assisted basic piano teaching in colleges and universities. Although digital piano classes in colleges and universities have been offered for many years, there are still various problems and confusions in the actual teaching for various reasons. To reflect the advantages of this new combined teaching mode, take piano solo teaching as an example, and use the traditional piano individual lessons for teaching arrangements. This research study will reflect on these problems and their causes, using the theory of wisdom education, which will help the healthy development of intelligent piano-assisted basic piano teaching. The exploration of countermeasures for intelligent piano-assisted basic piano teaching also helps to improve the efficiency of intelligent piano-assisted basic piano teaching at the same time [6]. The visualization of sound provides people with a new form of visual and auditory expression for the appreciation of auditory art. The new visual interpretation of auditory art adds to the expressiveness of auditory art, providing people with the dual beauty of the combination of visual and auditory senses, and the visualization of sound is conducive to people's deep understanding of the formal structure and emotional content of artworks. For example, in the live performance of the Czech Philharmonic Orchestra, the designer graphically designed the performance rhythmically and displayed it in 3D synchronously with the performance, bringing a visual feast to the audience.

This paper has an in-depth study of piano timbre. Through the piano pronunciation principle and the actual piano signal, the factors that affect the piano timbre are deeply studied. In addition, various existing audio analysis and synthesis techniques have been studied in depth. This paper studies common pitch recognition algorithms and fast pitch recognition algorithms for computer processing. The piano tone is reconstructed through the basic model of the piano mono signal. According to the factors that affect piano timbre, a timbre synthesis model with timbre parameters and editable timbre is designed. We design a synthesis method based on the timbre synthesis model, which can meet the requirements of timbre accuracy and real-time timbre synthesis by computer. We carried out the software and hardware design and equipment selection of the piano sound library generation system. In addition, the language used for software system development is C#. The system can realize the functions of piano audio acquisition, timbre editing, and timbre synthesis. It is experimentally verified that the timbre synthesized by using the existing timbre model is not realistic enough and has the characteristics of obvious electronic synthesized timbre, which cannot meet the requirements of digital piano and electric piano for high-quality timbre. By establishing an audio synthesis model, timbre editing can be realized, and the timbre can be synthesized to meet the requirements of piano sound realism. If the agent chooses from these actions, this action is called developing what is currently known about the value of the action. Conversely, if the agent chooses nongreedy actions, it is called temptation because it can improve the estimation of the value of nongreedy actions. Define appropriate timbre parameters to meet the dual requirements of timbre modification and timbre realism. Design a timbre synthesis method that enables rapid computer synthesis of timbres in real time while satisfying timbre realism and high accuracy. Design and implement a piano timbre library generation system to meet the common audio processing and timbre library synthesis functions. Several existing timbre synthesis models are introduced and simulated, but the synthesized timbres are not ideal; an audio synthesis model with editable timbres is proposed and the timbre synthesis steps are introduced; experiments show that timbre modification can be achieved by editing the timbre parameters of this model, and at the same time, the synthesized timbres meet the requirements of authenticity and high fidelity of piano timbre synthesis.

1.1. Status of Research. Pitch, timbre, intensity, and length are the four most important basic characteristics of music, and they lay the foundation for the entire piece of music. Pitch indicates the frequency of vibration of a sound; timbre refers to the different characteristics of a sound and represents different instruments; intensity, also known as loudness, indicates the size or strength of a sound; and length refers to the duration of a note from its beginning to its end [7]. In addition to the basic characteristics, there are complex characteristics like intervals, melody, rhythm, and overall characteristics like style, emotion, and category. Music features are an important basis for the study of music using computer technology, and the extraction of music features is an essential step for research in areas such as music retrieval and music signal processing. Music has many storage formats, and MIDI is a structured music data format that is widely used in research in the field of computer technology [8]. The total score of the work is excellent in the range of 90–100, good in the range of 80–89, average in the range of 70–79, qualified in the range of 60–69, and effort is needed in the range of 0–59. The design of piano fingering is closely related to the musical notes and piano keys, and the note information includes information such as pitch, pitch length, and pitch intensity, among which the pitch is associated with the piano keys; therefore, the pitch will be an

important musical feature for the automatic generation of piano fingering research in this paper [9]. As piano fingering is closely related to the key position of the fingers, and the key position information represents the pitch information of the notes in the concrete form of piano playing, if we know the key position of each note on the piano keyboard, then we can follow the fingering rules to complete the performance of the music [10]. Therefore, this paper will directly use the MIDI file of the music and extract the key information contained in it as the musical characteristics.

The positions and forms of hands and fingers are represented as HMM states, and the resulting sequences of executed notes are modelled as emissions associated with HMM transitions, and a search is performed using the Viterbi algorithm to find the most likely sequence of state transitions [11]. Zioga et al. proposed an HMM-based piano fingering annotation method for both hands that can handle both single notes as well as chords [12]. In most of these results, machine learning is used to automatically generate piano fingering sequences based on hidden Markov models, since piano fingering sequences are to be generated based on the note sequences of music [13]. The rule-based knowledge base system approach mainly refers to music composition relying on symbolic operations and rule constraints. The advantage of this type of approach is that it has accurate logical reasoning and every act made by the system can be interpreted. A genetic evolutionary algorithm is a method that imitates the process of the genetic evolution of species in biology, first constructing a suitable adaptation function, and then deciding the evolutionary direction of candidate chromosomes through the adaptation function to perform global optimization, which is considered a very effective method in the field of music generation [14]. Being able to imitate production works and even make innovative works proves the applicability of the design. The questionnaire was used to collect frontline teachers' attitudes and opinions on the design of interactive learning resources and the learning effects of students using interactive learning resources. Most teachers gave positive evaluations to the design and production of interactive learning resources. The main problems of neural network-based music generation are as follows: the main problem of neural network-based music generation is that the music features are more complex and the training method is too single, so the quality of the generated music is not guaranteed. However, the neural network-based music generation method has less human involvement and does not require professional knowledge of music theory, and once the model is trained, the generation speed is fast and music can be generated [15]. Therefore, in this paper, after studying the existing music generation techniques, we choose the neural network-based music generation method.

To make the model fully consider the real music composition process, this paper reviewed the information about music composition and interviewed music composers about the specific process of composing music. After further research, these abstract conceptual constraints are converted into computer language, and a chord progression reward mechanism and a music theory rule reward mechanism are constructed and introduced into the critic network constructed in this paper.

2. Analysis of Piano Tone Teaching in the Context of Artificial Intelligence Interaction

2.1. Artificial Intelligence Interaction Background Design. There are two directions in the music generation problem, monophonic music generation, and multitrack music generation. There is no good solution so far on how to represent the connection between different tracks with mathematical models in multitrack music generation, and multitrack music generation requires more stringent training data compared to monotrack. The existence of these problems causes the music tracks generated by multitrack music algorithms to be confusing, of low quality, and lacking in artistic aesthetics; therefore, the music generation in this paper is mainly aimed at monotrack music generation [16]. There are many fields involved in deep learning and many types of network models have emerged during the development process, convolutional neural networks (CNNs), RNNs, LSTMs, etc. Among the abovementioned network types, CNN is mainly used in the field of image processing to extract feature information in images utilizing convolutional kernel, and it can build multi-layer convolution to extract higher dimensional information features, which is beneficial to image recognition and classification. Hoch Reiter and Schmid Huber proposed the LSTM network framework, whose structure is shown in Figure 1. RNN adds three gates, namely, a forget gate, an input gate, and an output gate. Through the gate to control the information, and protect and control the state of the cells, the effect of learning long-term dependence on information is achieved.

In the implementation of interactive sound visualization, what determines the final visual representation of sound is the transformation patterns and association rules between sound and visual information, and the final visual representation presents a wide variety of variability due to different transformation patterns and rules. Conducive to image recognition and classification. RNN is mostly used in the field of natural language processing. Unlike the feedforward neural network, RNN introduces a self-loop structure, which is characterized by being able to deal with issues related to the input information. Randy and Ben define this way of transforming sound information to visual information through some attribute of association rules as "mapping," where different modes of transformation and association rules are different mappings. The term "mapping" is originally a mathematical term that refers to the unique correspondence between elements in a set of two elements [17]. In the mapping of sound visualization, sound information and visual information form two source bases for information matching, and some association rules are used to correspond to the information in the source bases and finally realize the transmission and transformation between audiovisual information.
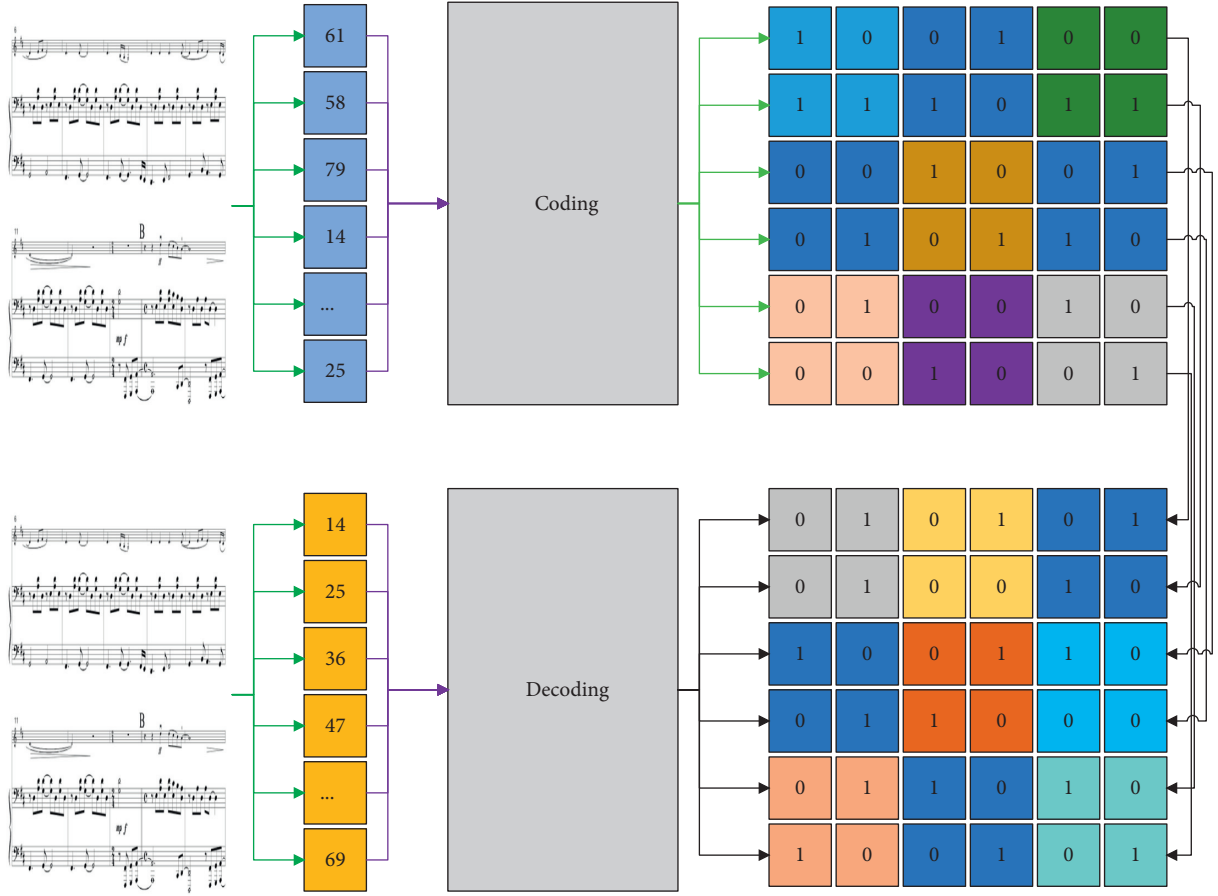
Figure 1: Generating a music framework.

It is different from mathematical mapping, and the "mapping" in this paper emphasizes some overall connection and transformation rules between audiovisual information from the macroscopic aspect, rather than referring to the unique correspondence of each element in it.

$$f_t = \sigma\left(W_f \cdot [h_t, X_t] - b_f\right). \tag{1}$$

The adaptation of learning resources to learners is manifested by the fact that learners can choose the learning content and the presentation of the learning content according to their own needs. The adaptation of learning resources to learners is manifested by giving immediate response and intelligent processing according to the dynamic information generated in the learning process of learners, such as providing different learning contents for learners according to the different answer situations of learners; recording students' learning; simulating problem situations in life, etc. And finally. realize the transmission and conversion between audiovisual information, which is not completely equivalent to mathematical mapping. The "mapping" in this article emphasizes a certain overall connection and conversion rule between audiovisual information from a macro perspective. Therefore, the support of learning resources for operational interaction can be summarized in five aspects: learner control, adaptivity, convenience, learning to monitor, and contextuality.

$$i_t = \sigma\left(W_i \cdot [h_t, X_{t-1}] - b_i\right), \tag{2}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, X_{t-1}] + b_o\right), \tag{3}$$

$$C_t = \coth\left(W_c \cdot [h_{t-1}, X_{t-1}] - b_c\right). \tag{4}$$

This stage is not an isolated existence, from aesthetic intuition to aesthetic understanding can produce relevant emotional and affective experience, it is the central outlet of various psychological content and form network structure in the aesthetic process, throughout the entire aesthetic process. When people are assimilated by the content and melody of sound work, they will integrate their body and mind into the sound art and produce the corresponding emotional experience, as in Gestalt psychology's mind-object isomorphism. This emotional experience unfolds along with the perception and understanding of the work and is an involuntary and intense psychological experience, which will be retained in the aesthetic memory for a long time even after the aesthetic process is over.

Sound is a type of wave and therefore has the common properties of waves. Frequency, amplitude, and period are three important physical quantities that can describe the properties of sound waves. Frequency is the number of periodic regular changes per second, and for the frequency of sound, it is the number of sound waves that produce

periodic vibrations in a second from the source. Amplitude is the maximum of the range of vibrations that a vibrating physical quantity can reach, and the amplitude of a sound determines the size of the sound. Sound waves exhibit periodic changes in their vibrational state during propagation, which describes the arrangement of individual repeated pressure changes of sound waves. The motion of the molecules of the medium is not visible to people, but according to the process of physical sound generation and physical properties, the conversion of the audiovisual mapping of physical sound can be achieved in three directions. Sound waves can be represented numerically and precisely by physical quantities such as frequency and amplitude, and through the perception of the extremely physical properties of the physical sound generation process, sound waves in physics are usually described and calculated mathematically in terms of the commonality of their generation principles, that is, a curved sonogram describing the periodic changes of sound energy at different moments experiencing different displacements utilizing mathematical number axes, as shown in Figure 2. The sonogram can accurately show the numerical variation of physical quantities of sound waves such as frequency and amplitude with time, which is important for the analysis of sound waves.

According to the physical nature of sound waves, a mapping transformation can be performed utilizing visual description, at which point the sound map becomes another form of sound existence, a visual feature of physical sound in human perception of sound, an external representation of sound that is based on objective knowledge of the physical characteristics of sound with unique certainty. Professional sound processing and editing software are based on this form of audiovisual conversion.

$$L(\theta_M) = \frac{1}{I}\sum_{i=1}^{I}(Y \ln Y - (1 + Y)\ln(1 + Y)). \qquad (5)$$

Software table technology uses algorithms to replace the hardware timbre library on the sound card and playback MIDI instrument sounds via CPU calculations in real time. The advantages of software table technology are that it produces high-quality sounds, is inexpensive, and is easily upgradeable [18]. Corresponding emotional experience is generated, such as the same shape of mind and object in Gestalt psychology, and psychological emotions will be generated with physical perception. However, the software table synthesis algorithms of different companies in the market are all trade secrets and cannot be obtained for free, and the synthesis results of software table synthesis algorithms vary from company to company, and their levels vary greatly. Therefore, an in-depth analysis of tone synthesis algorithms is needed.

$$P[S_{t+1} \mid S_t] = PP[S_{t+1} \mid S_1, \ldots, S_1]. \qquad (6)$$

There is a reason for the creation of new things, and humankind has never stopped searching and innovating in the field of art. Studies of the piano's development have shown that it is a product of the fusion of music and science
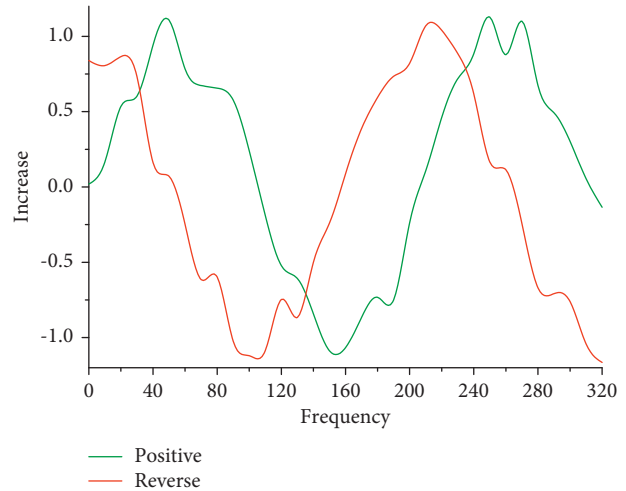

Figure 2: Sonogram.

and that it has become progressively more robust because of the improvement of science and technology. The piano is a keyboard instrument that produces sound by striking the strings through the power of the wrist, hand, and fingertips, so it also belongs to the category of percussion instruments, combining the attributes of a harmonious and dazzling instrument and a percussion instrument. But whatever changes this instrument underwent as it moved forward, it ultimately had a purpose around it, one to fill in the scraps of the age for the aural world with the finest sounds or combinations of different kinds of sounds, from which it gradually changed into aesthetic satisfaction.

$$\text{Notes Accuracy} = \frac{\sum_{m=1}^{M}\sum_{i=1}^{N} p(y_{im}, y_{im}^2)}{M + N}. \qquad (7)$$

Human language as instruction has two different forms of material bearing, namely, speech and writing. Speech belongs to the formal expression of language in the human body, and writing belongs to the symbolic form of the external presentation of language. The two belong to the expression of language in different spatial dimensions. Under the structured system of human linguistic knowledge, speech and text correspond to each other, and text is the visual representation of speech. Human beings have a special function of converting speech and words in real time and accurately, which is done according to human speech perception. Pure Tone collects the high-quality timbres of 9 famous European pianos, up to 7 layers of timbre performance, and 128 realistic timbres. The reading centre in the brain is responsible for linking the visual and auditory symbols of words according to the intentional requirements of the words, thus allowing the person to establish the audiovisual link of words, i.e., to hear the pronunciation and write the words. Usually, such recognition is uniquely deterministic within the same linguistic structural system. The audiovisual transformation between the two bearer forms of language is based on the accepted standard of decoding instructions for human linguistic knowledge, has a broad agreement in human society, and produces a visual commonality in people.

2.2. Experimental Analysis of Piano Tone Teaching. In the timbre composition analysis, we have concluded that all frequency components in the frequency domain work together to form a full and rich piano tone, and that overtone frequency components alone cannot be used for piano tone synthesis. The next step is to investigate the different effects of different frequency components on the formation of piano timbre [19]. In piano timbre, the fundamental and its overtone series play a major role in forming the timbre, and the timbre synthesized using only this frequency component can simulate the main timbral qualities of the piano. The frequencies of the telharmonic play an essential role in forming a full and rich piano tone, and the more interharmonic frequencies used in synthesis, the closer the synthesized tone is to the original piano tone. The part of the frequency spectrum whose frequencies are below the fundamental frequency forms the sound of a piano key struck now it is pressed. The intervals are divided into concordant and discordant intervals according to the auditory perception that the harmonic intervals give to the human ear. Every art form can bring people a dual aesthetic experience of physical and psychological. In the traditional sense, auditory art can only be experienced through auditory senses, and the form of experience is relatively simple. The interval's harmony is a characteristic that reflects the basic nature of the interval and is determined by the ratio of the vibrational frequencies of the two tones forming the interval. An interval that sounds pleasant and blends are called a concordant interval, while an interval that sounds harsher and does not blend is called a discordant interval. Consonant intervals include very fully consonant intervals, fully consonant intervals, and imperfectly consonant intervals. The relationship between the degree of interval harmony is shown in Figure 3.

From Figure 3, the time-domain waveforms of the three synthesized timbres differed among the three sets of experimental parameters, indicating that the parameters of the synthesis model can effectively change the time domain waveforms of the timbres. At the same time, the time-domain waveforms of the synthesized timbres are very similar in terms of the general characteristics of the waveform changes, and they all satisfy the three phases of piano articulation: onset, decay, and fade [20]. The frequency-domain waveform plots of the three synthesized timbres are different for the three sets of experimental parameters, indicating that the parameters of the synthesis model can effectively change the frequency domain waveforms of the timbres. Sound visualization provides a new form of expression that combines audiovisual art for people to appreciate auditory art. The new visual interpretation of auditory art adds the expressive power of auditory art, provides people with the dual beauty of audiovisual sensory, and visualizes sound. Performance is conducive to people's deep understanding of the form, structure, and content emotions of artistic works. At the same time, the frequency-domain waveforms of the three timbres have the same characteristics: the frequency distribution is clean and clear, the frequency amplitude is larger at the octave, and the rest of the frequency amplitude is smaller, and there are more

frequency components with smaller amplitude at the lower frequency band, which form the percussive sound of the piano now of pressing. Therefore, it can be judged that it belongs to the same category of piano tone.

The synthesized model timbre parameters can be set to change the time domain waveform and frequency domain waveform of the timbre and combined with the factors affecting the timbre, the timbre parameters must be changed to change the synthesized timbre; at the same time, the synthesized timbre retains the characteristics of the piano timbre in the time domain and frequency domain, in the time domain: it has three phases: onset, decay, and fade, and the sound rapidly decreases in the waveform and then continues to oscillate; in the frequency domain: the frequency distribution is clean. In the frequency domain, the frequency distribution is clean and clear, with larger frequency amplitude at the octave and smaller amplitude at the rest of the frequencies, and more frequency components of smaller amplitude at the low-frequency band that form the piano percussion. The synthesized timbre is a piano timbre, which satisfies the requirements of the piano timbre library generation system for timbre authenticity and high fidelity. Tone editing experiments can prove that using the audio synthesis touch pattern of editable tones in this paper for tone synthesis can change the synthesized tones by adjusting the tone parameters, and at the same time can meet the high-fidelity requirements of the piano tone generation system.

The main steps of piano tone library generation using tone library generation software include: source input, parameter editing, tone library generation, and the software displays audio/tone information and supports simulated playback of synthesized tones to assist the user in generating satisfactory tones and tone libraries. Explore the basic model of piano monophonic signal, reconstruct piano timbre through the basic model; design timbre parameters and an audio synthesis model of editable timbre according to the factors that affect piano timbre. The flowchart of tone library generation is shown in Figure 4.

The software supports two types of audio input methods: piano sound acquisition and import of existing audio files, which requires denoising of the audio after piano sound acquisition and audio file parsing to obtain audio data information. Then, the audio data spectrally transformed and the timbre characteristics are analysed; the timbre parameters are obtained through the human-computer interaction interface and the piano timbre is generated. If you are satisfied with the sound effect, you can save it as an audio file or directly generate and save the piano sound library in batch; if you are not satisfied with the sound effect, you can re-edit the sound parameters through the HMI to generate the sound.

With the combination of smart piano and traditional piano teaching, the teaching arrangement is divided into individual and group lessons crossed by piano solo teaching as an example. Six weeks as a systematic teaching cycle can be based on the intelligent feedback of the intelligent piano and the feedback of the manual teacher as experimental data, to do a summary analysis of the course progress, quality. The first and second weeks are in the form of group lessons on
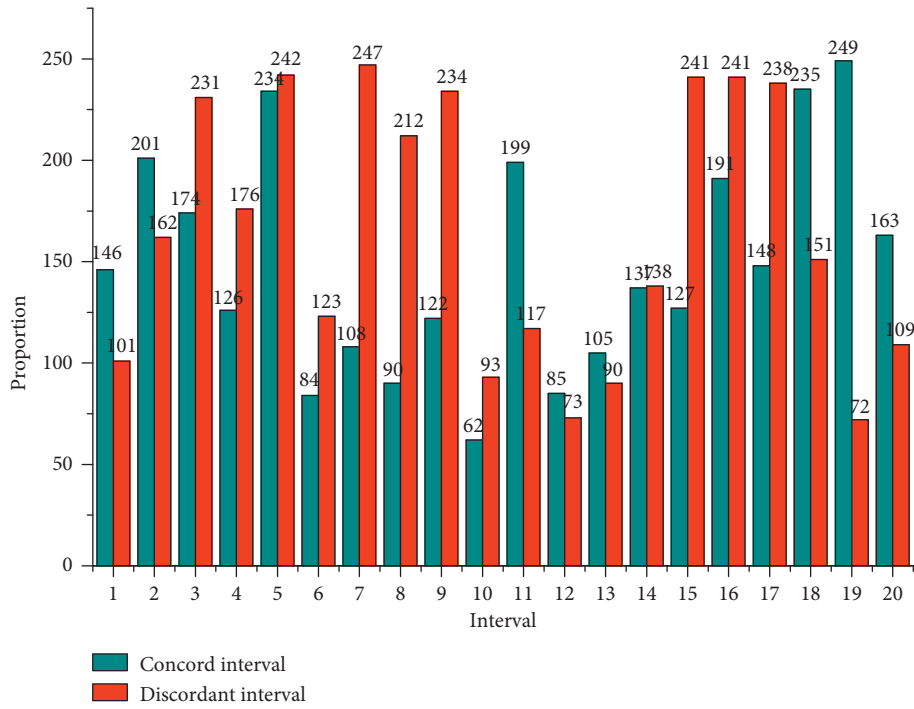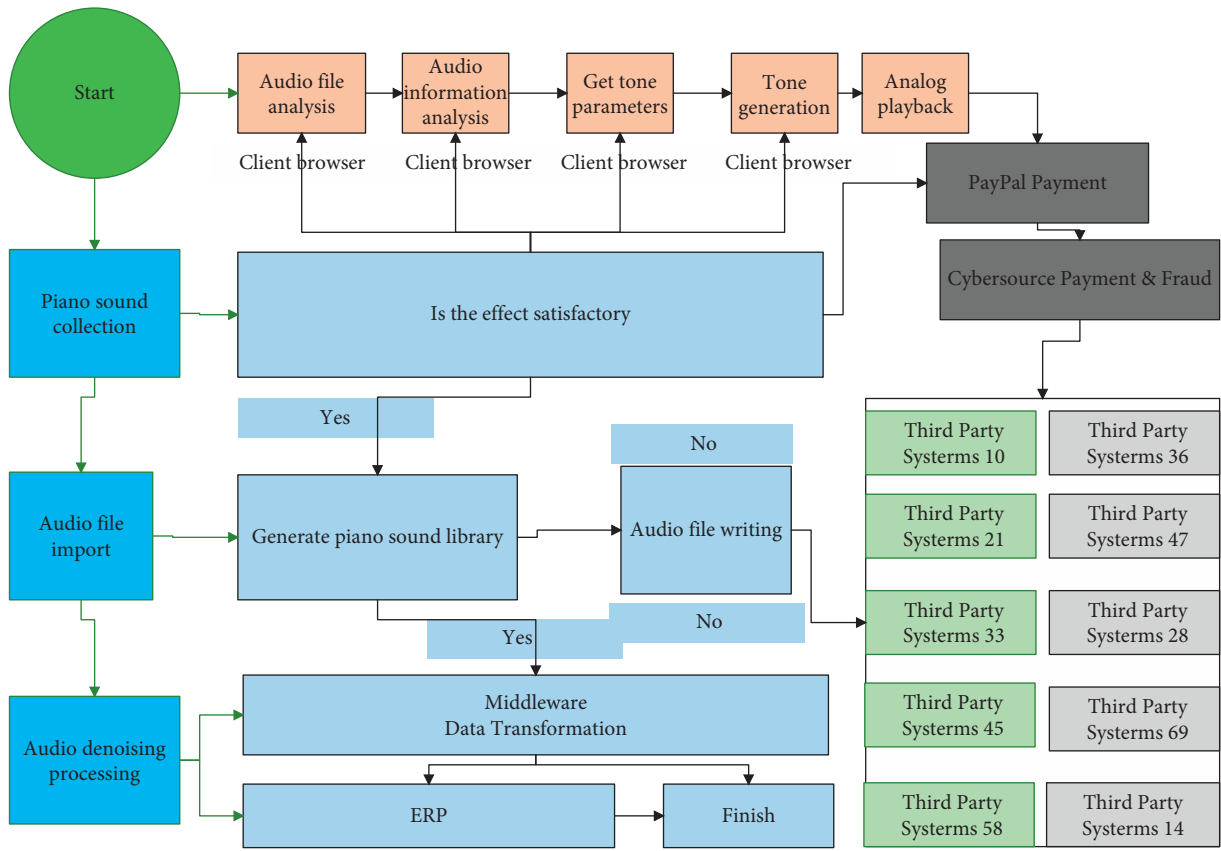
Figure 3: Interval concordance relationship.

Figure 4: Flowchart of tone library generation.

smart piano, the third and fourth weeks are in the form of individual lessons of smart piano, and the fifth and sixth weeks are in the form of individual lessons of a traditional piano. Taking the example of "Chopin's Narrative in G minor" as a case study, we started teaching from the first week to the sixth week, based on the combination of students' receptiveness, basic piano ability, and the difficulty level of "Chopin's Narrative in G minor," from the regular step-by-step teaching process the course is based on the combination of students' receptiveness, basic piano ability, and the difficulty level of the "Chopin's G minor." Each week, student's feedback is combined with the final assessment of student completion, thus demonstrating the advantages of this new combined teaching model. The system can realize the functions of piano audio collection, timbre editing, and timbre synthesis. It has been verified by experiments that the timbre synthesized by the existing timbre model is not real enough and has obvious characteristics of electronically synthesized timbre, which cannot meet the high-quality requirements of digital pianos and electric pianos. As an example, solo piano instruction is arranged in the form of a single traditional individual piano lesson [21]. The six weeks are used as a systematic teaching cycle, and the weekly lessons are summarized and analysed based on the combination of manual teacher feedback and intelligent feedback from the smart piano in the sixth week as experimental tracking data. The first six weeks were spent in the form of traditional individual piano lessons. Taking the "Chopin narrative in G minor" as a case study, the first week to the sixth week, based on the combination of students' receptiveness, basic piano ability, and the difficulty of the "Chopin narrative in G minor", each week combined with the data obtained from students' feedback, and finally checked students' completion. The final check of the students' completion is based on the data obtained from their feedback each week.

The two cases of the six weeks were compared in terms of different lesson formats and teaching methods, and the comparison shows that the combination of smart piano and traditional piano teaching is feasible and replicable. It is possible to try to incorporate "smart piano" into the existing piano teaching. Piano lessons with the "smart piano" application use both individual and group lessons, which are no longer mechanical and monotonous when combined with the traditional individual lessons. This new form of teaching is mainly based on intelligent technology, which can use pictures and audio to teach the students what they need to teach in oral and written form. Compared with the traditional teaching model, this is more stringent for students' learning quality, widens the path of students' knowledge acquisition, and improves teaching efficiency and quality.

3. Results and Analysis

3.1. Artificial Intelligence Interaction Performance Results. During the process of reinforcement learning, the intelligence will continuously estimate the value of states and actions, which is the basis for making action selection. When training reaches a mature stage, then the intelligence will mostly choose the action with the highest estimated value to perform, which will maximize the total gain and achieve our goal. However, at the beginning of training, the intelligence does not know which actions are really "good actions," so the problem of how to make action selection is called the action selection mechanism. The intelligence continuously estimates the value of an action so that at any given moment there is at least one action with the highest estimated value, and these actions are called greedy actions. If the intelligence chooses from these actions, this behaviour is called exploiting (exploit) the currently known knowledge about the value of the action. Conversely, if the intelligence chooses a nongreedy action, it is said to explore, because this improves the estimation of the value of the nongreedy action. "Exploitation" is correct for maximizing the desired benefit in the current moment, but "exploration" may maximize the overall benefit overall. And in an active choice, exploitation and probing cannot be performed simultaneously.

The average gain rate comparison statistics were obtained from the piano fingering generation experiments based on the constructed Q learning network using the key sequences of three music clips respectively, as shown in Figure 5. The average gain rate is the average of the ratio of the total gain to the optimal gain of a training ground, which reflects the superiority of the learning effect and the goodness of the result of the intelligence averaged over a certain range of training steps.

From Figure 5, when the number of training steps is relatively small, the average gain rate of each music fragment is unstable, the results are not satisfactory and fluctuate greatly, and as the number of training steps increases, the bits of intelligence learn to acquire better and better fingering rules, and the total gain gradually increases. The average gain rate stabilized earlier and later due to the different number of keys in each segment, and the segments with more keys require a larger number of training steps to ensure a better experimental result. The fingering sequences with larger yields were selected as the final output fingering sequence results in the training step interval where the average yield of each music fragment stabilized, respectively.

In terms of the cumulative score trend, the machine-generated fingering may be slightly lower than the professional fingering in the intermediate steps, but in general, the scoring trend of the machine-generated fingering for the three music clips is consistent with the professional fingering; in terms of the score rate trend, the scoring rate of the machine-generated fingering for clip 1 and clip 2 is consistent with the professional fingering, and the scoring rate of the machine-generated fingering for clip 3 shows some fluctuations. This indicates that there is still some room for improving the stability of the generated fingering results for music fragments with many notes, but for music fragments with a small number of keys, it can generate high-quality, stable fingering, as shown in Figure 6.

After the above experiments, the character-level music generation network Melodists constructed in this paper can accomplish the end-to-end music generation task and can learn the dependencies of notes on the time scale from the training data, but the Melodists network only uses pure LSTM
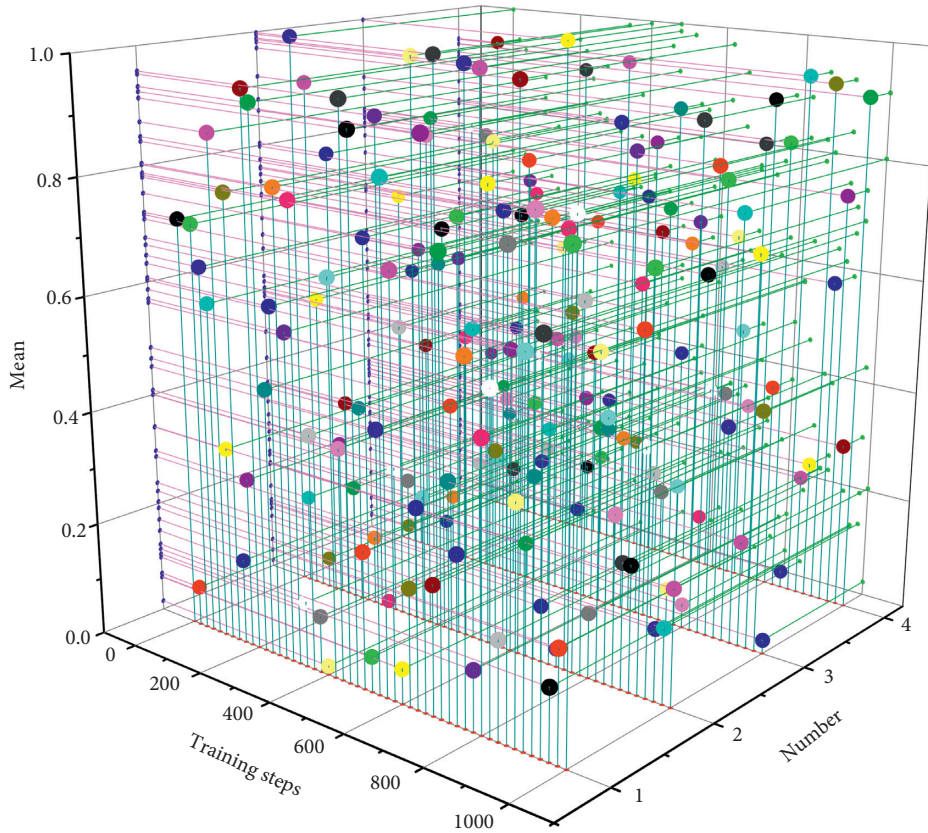
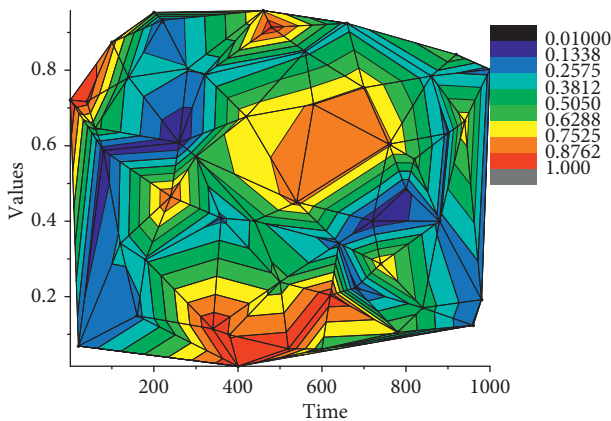Figure 5: Comparison of average yield across music segments.

Figure 6: Plot of generated music versus sample music and difference sound spectrum.

units and the structure is too simple that it cannot learn such complex data as music well. In this paper, we analyse many generation results and find that the generated music often contains many duplicate fragments, the generated music lacks structure, does not follow the rules of music theory, and lacks artistic aesthetics, and we also find that these problems exist in existing music algorithms. Therefore, this paper improves on the Melodists network constructed above and introduces a reinforcement learning algorithm to solve the above problems by combining deep learning and reinforcement learning.

3.2. Results of Teaching Piano Tone. To test whether students achieve a deeper level of conceptual interaction when using interactive learning resources for learning, the teacher should evaluate the mind maps drawn by students during the course, referring to the evaluation criteria for scholarly mind maps in terms of whether the mind maps meet the cognitive level of the students, the appropriateness of the choice of keywords, the connection of relationships between key objects, the presence of scientific errors, aesthetics, etc. The statistics of the evaluation levels of the student-drawn mind maps were obtained as shown in Figure 7.

Through the above evaluation of the students' mind maps, most of the students' mind maps were able to reach the level of C and above, and the mind maps were in line with the student's cognitive level, with appropriate keywords, accurate hierarchical relationships, and no scientific errors, but the aesthetic aspects of the mind maps were still lacking. Students were able to make correct connections between the new content learned in this lesson and the mind maps are drawn in the previous lesson, clarify the hierarchical relationships between concepts, and make correct connections between the new content and the content they had already learned in their minds in the previous lesson.

To test the effectiveness of students' learning using the interactive learning resources, and intelligent design work evaluation gauge is designed to evaluate the completion of students' classroom exercises in each lesson. Moreover, the multitrack music generation requires more stringent
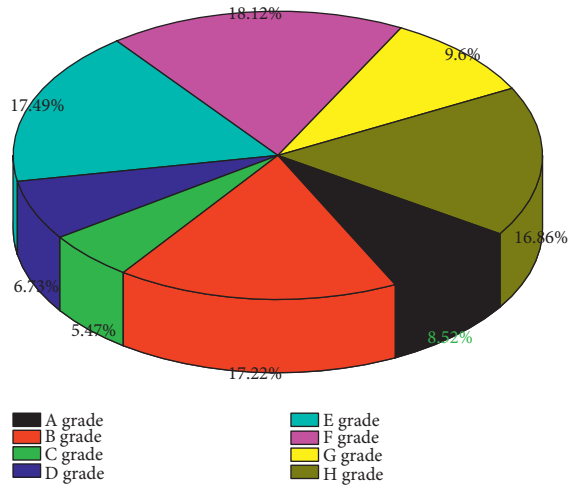
FIGURE 7: Thinking maps evaluation results statistics.

FIGURE 8: Comparison of the time-domain waveforms of the synthesized tones of the four sets of experimental parameters.

training data than the single-track music. The existence of these problems causes the confusion, low quality, and lack of artistic aesthetics of the music tracks generated by the multitrack music algorithm. The evaluation gauge evaluates the work in three dimensions: hardware assembly, pictorial programming, and work completion, with ten dimensions and a score of 10 points for each dimension. The total score of the work is excellent in the range of 90–100, good in the range of 80–89, fair in the range of 70–79, pass in the range of 60–69, and needs work in the range of 0–59. The work of 10 students within the AI club was evaluated according to the scale, resulting in the student work evaluation grade statistics shown in Figure 8.

The Smart Design Work Evaluation Scale evaluates students' classroom practice work, and the applicability of the design is demonstrated by the fact that students met the learning objectives after learning with the interactive learning resource and were able to imitate and produce work or even make innovative work. The questionnaire collected frontline teachers' attitudes and opinions about the design of the interactive learning resources and the learning effects of students' use of the interactive learning resources. Most of the teachers gave positive comments on the design and production of the interactive learning resources and believed that the use of the interactive learning resources could guide students to actively control their learning progress, independently assess their learning status, promote the independent construction of learners' knowledge, and realize the deep active interaction between learners. The use of interactive learning resources can guide students to actively control their learning progress, assess their learning status, promote the independent construction of learners' knowledge, and achieve deep active interaction between learners and learning resources.

Most surveyors say that smart pianos mostly used in group lessons and the equipment is old, and students rarely use smart pianos in their daily piano training. In the scenario of basic piano teaching, smart pianos are useful for technique teaching and psychological teaching; in the more complex and high-end teaching scenario involving
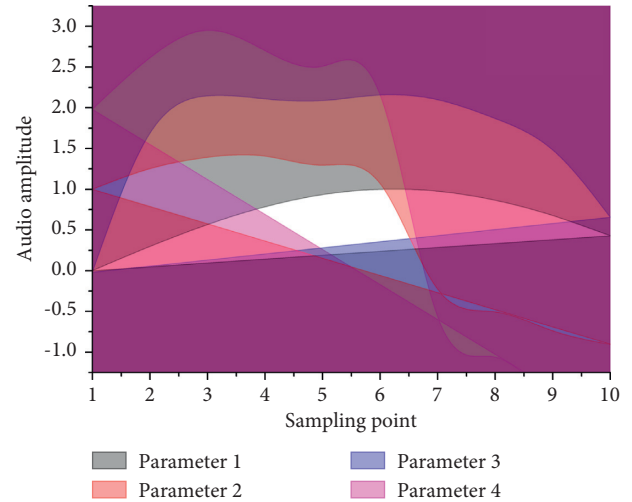
performance teaching, smart pianos are not suitable for this purpose. In the current educational environment, the use of intelligent pianos is still very limited. In the current situation of intelligent pianos to assist in basic piano teaching, the author first analyses the causes of the problem from the results of the survey and the literature, including the pressure of teachers to research, the uneven level of students, the lack of professional orientation for students, the lack of contact with intelligent piano courses and the lack of continuity in learning, and the lack of investment in hardware in schools.

4. Conclusion

In this paper, an audio synthesis model with editable timbre is proposed, in which the timbre parameters: maximum octave, amplitude modification coefficient, synthesis parameters, and modification type are set to realize the editable timbre. The synthesis and editing experiments show that the timbre synthesized by this model can meet the high-fidelity requirements of the piano timbre generation system. The model can be used to synthesize timbres to meet the flexibility and authenticity requirements of the piano timbre generation system. The existing piano timbre synthesis model is analysed and experimented with, and the difference between the synthesized timbre and the actual piano timbre is very large, which cannot meet the high-fidelity requirement of the sound source. In this paper, based on the multiplicative harmonic model of monophonic signal, combined with the variability of timbre characteristics, an audio synthesis model with editable timbre is proposed, and the editable timbre is realized by modifying each timbre parameter of the model. The experimental results show that the piano timbre synthesized by this model can meet the system's requirements for flexibility and realism of timbre library generation. The system software includes audio file reading and writing touch block, audio information analysis module, timbre parameter acquisition module, timbre synthesis module, and simulated performance module.

Compared with the traditional methods, the use of this system to generate piano timbre libraries is inexpensive and simple to operate, and it meets the demand for authenticity and personalization of sound sources for digital pianos and smart pianos.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] A. Zamm, S. Debener, I. Konvalinka, N. Sebanz, and G. Knoblich, "The sound of silence: an EEG study of how musicians time pauses in individual and joint music performance," *Social Cognitive and Affective Neuroscience*, vol. 16, no. 1-2, pp. 31–42, 2021.

[2] D. Johnson, D. Damian, and G. Tzanetakis, "Detecting hand posture in piano playing using depth data," *Computer Music Journal*, vol. 43, no. 1, pp. 59–78, 2020.

[3] S. Canazza and G. D. Poli, "Four decades of music research, creation, and education at padua's centro di sonologia computazionale," *Computer Music Journal*, vol. 43, no. 4, pp. 58–80, 2020.

[4] M. Mueller, B. McFee, and K. Kinnaird, "Interactive learning of signal processing through music: making fourier analysis concrete for students," *IEEE Signal Processing Magazine*, vol. 38, no. 3, pp. 73–84, 2021.

[5] J.-X. He, L. Zhou, Z.-T. Liu, and X.-Y. Hu, "Digital empirical research of influencing factors of musical emotion classification based on pleasure-arousal musical emotion fuzzy model," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 24, no. 7, pp. 872–881, 2020.

[6] N. Siphocly, A.-B. Salem, and E.-S. El-Horabty, "Applications of computational intelligence in computer music composition," *International Journal of Intelligent and Cooperative Information Systems*, vol. 21, no. 1, pp. 59–67, 2021.

[7] N. N. J. Siphocly, E. S. M. El-Horbaty, and A. B. M. Salem, "Top 10 artificial intelligence algorithms in computer music composition," *International Journal of Computing and Digital Systems*, vol. 10, no. 01, pp. 373–394, 2021.

[8] A. Pošćić and G. Kreković, "On the human role in generative art: a case study of AI-driven live coding," *Journal of Science and Technology of the Arts*, vol. 12, no. 3, pp. 45–62, 2020.

[9] M. Lesaffre and M. Leman, "Integrative research in art and science: a framework for proactive humanities," *Critical Arts*, vol. 34, no. 5, pp. 39–54, 2020.

[10] L. Gabrielli, S. Tomassetti, and S. Squartini, "Timbre equalization of wind instruments physical models using a random iterative search algorithm," *Journal of the Audio Engineering Society*, vol. 68, no. 5, pp. 364–376, 2020.

[11] Y. Deldjoo, M. Schedl, P. Cremonesi, and G. Pasi, "Recommender systems leveraging multimedia content," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–38, 2020.

[12] I. Zioga, P. M. C. Harrison, M. T. Pearce, J. Bhattacharya, and C. D. B. Luft, "Auditory but not audiovisual cues lead to higher neural sensitivity to the statistical regularities of an unfamiliar musical style," *Journal of Cognitive Neuroscience*, vol. 32, no. 12, pp. 2241–2259, 2020.

[13] C. Jin, T. Wang, S. Liu et al., "A transformer-based model for multi-track music generation," *International Journal of Multimedia Data Engineering & Management*, vol. 11, no. 3, pp. 36–54, 2020.

[14] G. Emerson and H. Egermann, "Exploring the motivations for building new digital musical instruments," *Musicae Scientiae*, vol. 24, no. 3, pp. 313–329, 2020.

[15] D. Ihde, "A finnish turn: digital and synthesiser musical instruments," *Journal of New Music Research*, vol. 50, no. 2, pp. 165–174, 2021.

[16] H. B. Lima, C. G. D. Santos, and B. S. Meiguins, "A survey of music visualization techniques," *ACM Computing Surveys*, vol. 54, no. 7, pp. 1–29, 2021.

[17] M. C. Lee, S. Y. Chiang, S. C. Yeh, and T. F. Wen, "Study on emotion recognition and companion Chatbot using deep neural network," *Multimedia Tools and Applications*, vol. 79, no. 27, pp. 19629–19657, 2020.

[18] D. Worrall, "Computational designing of sonic morphologies," *Organised Sound*, vol. 25, no. 1, pp. 15–24, 2020.

[19] M. Müller, "An educational guide through the FMP notebooks for teaching and learning fundamentals of music processing," *Signals*, vol. 2, no. 2, pp. 245–285, 2021.

[20] S. Oore, I. Simon, S. Dieleman, D. Eck, and K. Simonyan, "This time with feeling: learning expressive musical performance," *Neural Computing & Applications*, vol. 32, no. 4, pp. 955–967, 2020.

[21] E. Méchoulan and D. F. Bell, "Are sounds sound? for an enthusiastic study of sound studies," *Substance*, vol. 49, no. 2, pp. 3–29, 2020.